GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 23, 2006, 12:00:35 ; Search time 187 Seconds
(without alignments)
636.747 Million cell updates/sec

Title: US-10-620-562-2
Perfect score: 1426
Sequence: 1 MAPRALPGSAVLAAAVFVGG.....ETVNGEVPATPVKRERSGTE 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1426	100.0	271	5	AAE20962	Aae20962 Human TR2
2	1426	100.0	271	7	ADI21099	Adi21099 Novel hum
3	1421	99.6	271	4	AAB88406	Aab88406 Human mem
4	1421	99.6	271	9	ADY63175	Ady63175 Human clo
5	811	56.9	195	4	AAM23972	Aam23972 Human EST
6	811	56.9	195	7	ADI21579	Adi21579 Novel hum
7	477.5	33.5	103	2	AAW88579	Aaw88579 Secreted
8	477.5	33.5	103	4	ABB50346	Abb50346 Human sec
9	477.5	33.5	103	6	ABO44603	Abo44603 Novel hum
10	477.5	33.5	103	7	ABO26083	Abo26083 Human pro
11	332	23.3	181	4	ABB50685	Abb50685 Human sec
12	332	23.3	181	5	AAE20965	Aae20965 Human TR2
13	332	23.3	181	6	ABO44942	Abo44942 Novel hum
14	332	23.3	181	7	ABO26422	Abo26422 Protein a
15	310	21.7	382	8	ADP29453	Adp29453 Human sec
16	310	21.7	382	8	ADP29451	Adp29451 Human sec
17	307	21.5	408	2	AAY41111	Aay41111 Human TAN
18	307	21.5	430	2	AAY41110	Aay41110 Human TAN
19	307	21.5	430	3	AAY70785	Aay70785 Human tum
20	307	21.5	430	4	AAE05517	Aae05517 Human tum
21	307	21.5	430	4	AAU09901	Aau09901 Human tum
22	307	21.5	430	5	ABG31739	Abg31739 Human TR4
23	307	21.5	430	6	ADA83983	Ada83983 Human TNF
24	307	21.5	430	8	ABO84934	Abo84934 Human can
25	307	21.5	430	8	ABM81290	Abm81290 Tumour-as
26	307	21.5	430	9	ADZ13510	Adz13510 Human can
27	307	21.5	430	9	ADZ13516	Adz13516 Human can
28	307	21.5	457	9	ADZ13508	Adz13508 Human can
29	307	21.5	878	4	AAE02358	Aae02358 Human tum
30	296	20.8	429	7	ADC52548	Adc52548 human den
31	296	20.8	430	4	AAB95627	Aab95627 Human pro
32	296	20.8	430	5	ABG31740	Abg31740 Human TR4
33	296	20.8	430	6	ABP70909	Abp70909 Human obe
34	295.5	20.7	317	9	ADZ13501	Adz13501 Murine ca
35	295.5	20.7	359	8	ABO84931	Abo84931 Murine ca
36	295.5	20.7	429	9	ADZ13503	Adz13503 Murine ca
37	295.5	20.7	436	4	AAE05518	Aae05518 Mouse tum
38	284	19.9	303	4	AAM93499	Aam93499 Human pol
39	284	19.9	303	8	ADL31172	Adl31172 Human pro
40	281.5	19.7	245	6	ABR43195	Abr43195 Human REM
41	281	19.7	294	4	AAU87162	Aau87162 Novel cen
42	281	19.7	294	8	ADI54477	Adi54477 Novel hum
43	278.5	19.5	194	5	AAE20963	Aae20963 Human TR2
44	278	19.5	339	8	ADP29463	Adp29463 Human sec
45	240	16.8	266	8	ADP29422	Adp29422 Human sec

ALIGNMENTS

RESULT 1
AAE20962
ID AAE20962 standard; protein; 271 AA.
XX
AC AAE20962;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human TR21 receptor protein.
XX
KW Human; tumour necrosis factor receptor; TNF; TR21; TR22; immune disorder;
KW autoimmune haemolytic anaemia; rheumatoid arthritis; Addison's disease;
KW allergy; cancer; ulcerative colitis; cardiovascular disorder; epilepsy;
KW myocardial ischaemia; wound healing; neurological disease; infection;
KW cerebral anoxia; gene therapy; immunosuppressive; antiproliferative;
KW cytostatic; cardiant; vasotropic; neuroprotective; antibacterial;
KW virucide; fungicide; nootropic; ophthalmological.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .29
FT /label= Signal_peptide
FT Region 29. .56
FT /label= Immunogenic_epitope
FT Protein 30. .271
FT /label= Mature_TR21_protein
FT Region 81. .87
FT /label= Immunogenic_epitope
FT Region 109. .116
FT /label= Immunogenic_epitope
FT Region 142. .149
FT /label= Immunogenic_epitope
FT Region 168. .173
FT /label= Immunogenic_epitope
FT Region 192. .199
FT /label= Immunogenic_epitope
FT Region 204. .218
FT /label= Immunogenic_epitope
FT Region 233. .238
FT /label= Immunogenic_epitope
FT Region 262. .271
FT /label= Immunogenic_epitope
XX
PN WO200207762-A1.
XX
PD 31-JAN-2002.

XX
PF 23-JUL-2001; 2001WO-US023124.
XX
PR 24-JUL-2000; 2000US-0220116P.
PR 27-JUL-2000; 2000US-0221143P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Zeng Z, Ruben SM, Rosen CA;
XX
DR WPI; 2002-171948/22.
DR N-PSDB; AAD33366.
XX
PT Two novel proteins, TR21 and TR22, which are members of the tumor
PT necrosis factor receptor, useful for the diagnosis and treatment of
PT immune disorders, cancer, cardiovascular disorders.
XX
PS Claim 11; Fig 1; 248pp; English.
XX
CC The present invention relates to novel human tumour necrosis factor (TNF)
CC receptors, TR21 and TR22 and polynucleotides encoding them. Sequences of
CC the invention are useful in the diagnosis, treatment and prevention of
CC cancers (e.g., cancers of the adrenal gland, bone, urogenital or bone
CC marrow, in particular breast and ovarian cancer), immune disorders (e.g.,
CC autoimmune haemolytic anaemia, rheumatoid arthritis, allergies, Addison's
CC disease, ulcerative colitis), cardiovascular disorders (e.g., myocardial
CC ischaemias), wound healing, neurological diseases (e.g., cerebral anoxia,
CC epilepsy) and infectious diseases such as viral, bacterial, fungal and
CC parasitic infections. They are also useful in gene therapy. The present
CC sequence is human TR21 protein
XX
SQ Sequence 271 AA;

Query Match 100.0%; Score 1426; DB 5; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.1e-120;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180

Qy        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240

Qy        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
              |||||||||||||||||||||||||||||||
Db        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
```

RESULT 2
ADI21099
ID ADI21099 standard; protein; 271 AA.
XX
AC ADI21099;
XX
DT 15-APR-2004 (first entry)
XX
DE Novel human protein #74.
XX
KW forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
XX
OS Homo sapiens.
XX
PN WO2003025148-A2.
XX
PD 27-MAR-2003.
XX
PF 19-SEP-2002; 2002WO-US029964.
XX
PR 19-SEP-2001; 2001US-0323739P.
PR 13-SEP-2002; 2002US-00323739.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI Haley-Vicente D;
XX
DR WPI; 2003-354603/33.
DR N-PSDB; ADI21815.
XX
PT New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.
XX
PS Claim 20; SEQ ID NO 350; 156pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers,
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents the amino acid sequence of a novel human
CC protein.
XX
SQ Sequence 271 AA;

Query Match 100.0%; Score 1426; DB 7; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.1e-120;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180

Qy        181 HTVGGVVERDVCHRCRHKRWHPIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240

Qy        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
              |||||||||||||||||||||||||||||||
Db        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
```

RESULT 3

```
AAB88406
ID   AAB88406 standard; protein; 271 AA.
XX
AC   AAB88406;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Human membrane or secretory protein clone PSEC0162.
XX
KW   Human; secretory protein; membrane protein; vaccine; gene therapy;
KW   rheumatoid arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   EP1067182-A2.
XX
PD   10-JAN-2001.
XX
PF   07-JUL-2000; 2000EP-00114090.
XX
PR   08-JUL-1999;   99JP-00194179.
PR   11-JAN-2000; 2000JP-00118775.
PR   02-MAY-2000; 2000JP-00183766.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Kawai Y,  Sugiyama T,  Hayashi K;
XX
DR   WPI; 2001-093989/11.
DR   N-PSDB; AAF93833.
XX
PT   Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT   gene therapy or as candidate target molecules in drug development.
XX
PS   Claim 1; SEQ ID NO 180; 609pp + Sequence Listing; English.
XX
CC   This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC   which encode human secretory or membrane proteins represented by AAB88317
CC   - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and
CC   AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the
CC   invention. The invention also includes methods for the production of
CC   antibodies directed against the proteins, and cDNA sequences, which can
CC   be used in vaccines. The polynucleotide sequences can be used in gene
CC   therapy. The polynucleotide sequences and the proteins they encode may be
CC   used in the prevention, treatment and diagnosis of diseases associated
CC   with inappropriate secretory protein/membrane protein expression. The
CC   nucleic acids and complementary sequences may also be used as DNA probes
CC   in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect
CC   and quantitate the presence of similar nucleic acid sequences in samples.
CC   They may also be used to study the expression and function of secretory
CC   proteins/membrane polypeptides and their role in metabolism. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   against them and in assays to identify modulators (agonists and
CC   antagonists) of expression and activity. The antibodies and antagonists
CC   may also be used as therapeutic agents to down regulate expression and
CC   activity. The antibodies may also be used as diagnostic agents for
CC   detecting the presence of the polypeptides in samples (e.g. by enzyme
CC   linked immunosorbant assay (ELISA). Examples of diseases which may be
CC   treated include rheumatoid arthritis and diabetes
XX
SQ   Sequence 271 AA;

  Query Match             99.6%;  Score 1421;  DB 4;  Length 271;
  Best Local Similarity   99.6%;  Pred. No. 3.1e-120;
  Matches  270;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRALHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180

Qy        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240

Qy        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
              |||||||||||||||||||||||||||||||
Db        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271


RESULT 4
ADY63175
ID   ADY63175 standard; protein; 271 AA.
XX
AC   ADY63175;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Human clone PSEC0162 protein, SEQ ID 180.
XX
KW   Gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1514933-A1.
XX
PD   16-MAR-2005.
XX
PF   07-JUL-2000; 2004EP-00027228.
XX
PR   08-JUL-1999;   99JP-00194179.
PR   11-JAN-2000; 2000JP-00118775.
PR   02-MAY-2000; 2000JP-00183766.
PR   07-JUL-2000; 2000EP-00114090.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Kawai Y,  Sugiyama T,  Hayashi K;
XX
DR   WPI; 2005-203865/22.
DR   N-PSDB; ADY63174.
XX
PT   Novel isolated polynucleotide encoding human secretary proteins or
PT   membrane proteins, useful for examination and diagnosis of abnormality of
PT   human secretary proteins.
XX
PS   Disclosure; SEQ ID NO 180; 1240pp; English.
XX
CC   The present invention relates to novel human secretory proteins or
CC   membrane proteins, and their coding sequences. The present sequence is
CC   one such protein sequence. The coding sequences of the invention are
CC   useful for examination and diagnosis of abnormality of the human
CC   secretary proteins and in gene therapy methods. The coding sequences and
CC   proteins are useful as candidates for medicines or as target molecules
CC   for developing medicines. Antibodies against the proteins of the
CC   invention are useful for treating diseases that are associated with the
CC   proteins. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained from sequence information
CC   supplied by the European Patent Office.
XX
SQ   Sequence 271 AA;

  Query Match             99.6%;  Score 1421;  DB 9;  Length 271;
  Best Local Similarity   99.6%;  Pred. No. 3.1e-120;
  Matches  270;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRALHSRTETTPSPSNDTGNGHPEYIA 60
```

```
Qy         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180

Qy        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240

Qy        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
              ||||||||||||||||||||||||||||||
Db        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
```

RESULT 5
AAM23972
ID AAM23972 standard; protein; 195 AA.
XX
AC AAM23972;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST encoded protein SEQ ID NO: 1497.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
KW gene therapy; nutrition.
XX
OS Homo sapiens.
XX
PN WO200154477-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US002687.
XX
PR 25-JAN-2000; 2000US-00491404.
PR 17-JUL-2000; 2000US-00617746.
PR 03-AUG-2000; 2000US-00631451.
PR 15-SEP-2000; 2000US-00663870.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
DR WPI; 2001-476164/51.
DR N-PSDB; AAH98631.
XX
PT Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use.
XX
PS Claim 20; Page 1037-1038; 1275pp; English.
XX
CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention
XX
SQ Sequence 195 AA;

Query Match 56.9%; Score 811; DB 4; Length 195;
Best Local Similarity 99.4%; Pred. No. 4.3e-65;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSP 158
              |||||||||||||||||||||||||||||||||||||| |
Db        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGEP 158
```

RESULT 6
ADI21579
ID ADI21579 standard; protein; 195 AA.
XX
AC ADI21579;
XX
DT 15-APR-2004 (first entry)
XX
DE Novel human polypeptide #58.
XX
KW forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
XX
OS Homo sapiens.
XX
PN WO2003025148-A2.
XX
PD 27-MAR-2003.
XX
PF 19-SEP-2002; 2002WO-US029964.
XX
PR 19-SEP-2001; 2001US-0323739P.
PR 13-SEP-2002; 2002US-00323739.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI Haley-Vicente D;
XX
DR WPI; 2003-354603/33.
DR N-PSDB; ADI21359.
XX
PT New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.
XX
PS Example 3; SEQ ID NO 830; 156pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers,
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents the amino acid sequence of a novel human
CC polypeptide.

```
XX
SQ   Sequence 195 AA;

  Query Match             56.9%;  Score 811;  DB 7;  Length 195;
  Best Local Similarity   99.4%;  Pred. No. 4.3e-65;
  Matches  157;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSP 158
              |||||||||||||||||||||||||||||||| |
Db        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGEP 158


RESULT 7
AAW88579
ID   AAW88579 standard; protein; 103 AA.
XX
AC   AAW88579;
XX
DT   01-MAR-1999  (first entry)
XX
DE   Secreted protein encoded by gene 46 clone HCFMV39.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9854963-A2.
XX
PD   10-DEC-1998.
XX
PF   04-JUN-1998;   98WO-US011422.
XX
PR   06-JUN-1997;   97US-0048875P.
PR   06-JUN-1997;   97US-0048876P.
PR   06-JUN-1997;   97US-0048877P.
PR   06-JUN-1997;   97US-0048878P.
PR   06-JUN-1997;   97US-0048880P.
PR   06-JUN-1997;   97US-0048881P.
PR   06-JUN-1997;   97US-0048882P.
PR   06-JUN-1997;   97US-0048883P.
PR   06-JUN-1997;   97US-0048884P.
PR   06-JUN-1997;   97US-0048885P.
PR   06-JUN-1997;   97US-0048892P.
PR   06-JUN-1997;   97US-0048893P.
PR   06-JUN-1997;   97US-0048894P.
PR   06-JUN-1997;   97US-0048895P.
PR   06-JUN-1997;   97US-0048896P.
PR   06-JUN-1997;   97US-0048897P.
PR   06-JUN-1997;   97US-0048898P.
PR   06-JUN-1997;   97US-0048899P.
PR   06-JUN-1997;   97US-0048900P.
PR   06-JUN-1997;   97US-0048901P.
PR   06-JUN-1997;   97US-0048915P.
PR   06-JUN-1997;   97US-0048916P.
PR   06-JUN-1997;   97US-0048917P.
PR   06-JUN-1997;   97US-0048949P.
PR   06-JUN-1997;   97US-0048962P.
PR   06-JUN-1997;   97US-0048963P.
PR   06-JUN-1997;   97US-0048964P.
PR   06-JUN-1997;   97US-0048970P.
PR   06-JUN-1997;   97US-0048971P.
PR   06-JUN-1997;   97US-0048972P.
PR   06-JUN-1997;   97US-0048974P.
PR   06-JUN-1997;   97US-0049019P.
PR   06-JUN-1997;   97US-0049020P.
PR   06-JUN-1997;   97US-0049373P.
PR   06-JUN-1997;   97US-0049374P.
PR   06-JUN-1997;   97US-0049375P.
PR   05-SEP-1997;   97US-0057584P.
PR   05-SEP-1997;   97US-0057627P.
PR   05-SEP-1997;   97US-0057628P.
PR   05-SEP-1997;   97US-0057629P.
PR   05-SEP-1997;   97US-0057634P.
PR   05-SEP-1997;   97US-0057635P.
PR   05-SEP-1997;   97US-0057642P.
PR   05-SEP-1997;   97US-0057643P.
PR   05-SEP-1997;   97US-0057644P.
PR   05-SEP-1997;   97US-0057645P.
PR   05-SEP-1997;   97US-0057646P.
PR   05-SEP-1997;   97US-0057647P.
PR   05-SEP-1997;   97US-0057648P.
PR   05-SEP-1997;   97US-0057649P.
PR   05-SEP-1997;   97US-0057650P.
PR   05-SEP-1997;   97US-0057651P.
PR   05-SEP-1997;   97US-0057654P.
PR   05-SEP-1997;   97US-0057661P.
PR   05-SEP-1997;   97US-0057662P.
PR   05-SEP-1997;   97US-0057666P.
PR   05-SEP-1997;   97US-0057667P.
PR   05-SEP-1997;   97US-0057668P.
PR   05-SEP-1997;   97US-0057760P.
PR   05-SEP-1997;   97US-0057761P.
PR   05-SEP-1997;   97US-0057762P.
PR   05-SEP-1997;   97US-0057763P.
PR   05-SEP-1997;   97US-0057764P.
PR   05-SEP-1997;   97US-0057765P.
PR   05-SEP-1997;   97US-0057769P.
PR   05-SEP-1997;   97US-0057770P.
PR   05-SEP-1997;   97US-0057771P.
PR   05-SEP-1997;   97US-0057774P.
PR   05-SEP-1997;   97US-0057775P.
PR   05-SEP-1997;   97US-0057776P.
PR   05-SEP-1997;   97US-0057777P.
PR   05-SEP-1997;   97US-0057778P.
PR   18-DEC-1997;   97US-0070923P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Young P,  Greene JM,  Ferrie AM,  Ruben SM,  Rosen CA,  Hu J;
PI   Olsen HS,  Ebner R,  Brewer LA,  Moore PA,  Shi Y,  Florence C;
PI   Florence K,  Lafleur DW,  Ni J,  Fan P,  Wei Y,  Fischer CL,  Soppet DR;
PI   Li Y,  Zeng Z,  Kyaw H,  Yu G,  Feng P,  Dillon PJ,  Endress GA;
PI   Carter KC;
XX
DR   WPI; 1999-059865/05.
DR   N-PSDB; AAV84456.
XX
PT   New isolated human genes and the secreted polypeptides they encode -
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders.
XX
PS   Claim 11; Page 511; 772pp; English.
XX
CC   The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
CC   encoding human secreted proteins (AAW88534 to AAW88756). The secreted
CC   protein gene sequences are deposited with the ATCC under deposit numbers
CC   ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC   209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC   cells comprising recombinant vectors containing the nucleic acid
CC   sequences are used for the recombinant production of the secreted
CC   proteins. The polynucleotide and amino acid sequences are useful for are
```

```
CC   useful for preventing, treating or ameliorating medical conditions e.g.
CC   by protein or gene therapy. Pathological conditions can be also diagnosed
CC   by determining the amount of the new polypeptides in a sample or by
CC   determining the presence of mutations in the new polynucleotides.
CC   Specific uses are described for each of the polynucleotides, based on
CC   which tissues they are most highly expressed in, and include developing
CC   products for the diagnosis or treatment of cancer, neurodegenerative
CC   disorders, developmental abnormalities and foetal deficiencies, blood
CC   disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC   diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC   ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC   restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC   such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC   or thymus, digestive/endocrine disorders, infections and AIDS. The
CC   polypeptides are also useful for identifying their binding partners. The
CC   present sequence represents human secreted protein (see descriptor line
CC   for gene number and clone identification)
XX
SQ   Sequence 103 AA;

  Query Match             33.5%;  Score 477.5;  DB 2;  Length 103;
  Best Local Similarity   95.0%;  Pred. No. 3.3e-35;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||     ||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101


RESULT 8
ABB50346
ID   ABB50346 standard; protein; 103 AA.
XX
AC   ABB50346;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Human secreted protein encoded by gene 46 SEQ ID NO:294.
XX
KW   Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
KW   dermatological; immunosuppressive; antiinflammatory; immunostimulant;
KW   cytostatic; cardiant; vascular; anti-angiogenic; ophthalmological;
KW   neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnerary;
KW   antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW   multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW   human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW   Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis;
KW   Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder;
KW   corneal graft neovascularisation; diabetic retinopathy; regeneration;
KW   neurological disorder; Huntington's chorea; Alzheimer's disease;
KW   Parkinson's disease; infectious disease.
XX
OS   Homo sapiens.
XX
PN   WO200162891-A2.
XX
PD   30-AUG-2001.
XX
PF   21-FEB-2001; 2001WO-US005614.
XX
PR   24-FEB-2000; 2000US-0184836P.
PR   29-MAR-2000; 2000US-0193170P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J,  Ebner R,  Lafleur DW,  Moore PA,  Olsen HS,  Rosen CA;
PI   Ruben SM,  Soppet DR,  Young PE,  Shi Y,  Florence KA,  Wei Y;
PI   Florence C,  Hu J,  Li Y,  Kyaw H,  Fischer CL,  Ferrie AM,  Fan P;
PI   Feng P,  Endress GA,  Dillon PJ,  Carter KC,  Brewer LA,  Yu G,  Zeng Z;
PI   Greene JM;
XX
DR   WPI; 2001-625724/72.
DR   N-PSDB; ABA83239.
XX
PT   Nucleic acids encoding 207 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease
PT   and diabetic retinopathy.
XX
PS   Claim 11; Page 1103-1104; 1533pp; English.
XX
CC   ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
CC   proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
CC   activities based on the tissues and cells the genes are expressed in.
CC   Example of these activities include: immunomodulatory; antisclerotic;
CC   dermatological; immunosuppressive; antiinflammatory; immunostimulant;
CC   anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;
CC   neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;
CC   antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used
CC   in gene therapy and vaccine production. (I) and (II) can be used in the
CC   prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC   sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC   (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC   Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
CC   Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC   disorders (e.g. corneal graft neovascularisation and diabetic
CC   retinopathy), neurological disorders (e.g. Huntington's chorea,
CC   Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC   for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
CC   ABA83193 and ABB50300 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 103 AA;

  Query Match             33.5%;  Score 477.5;  DB 4;  Length 103;
  Best Local Similarity   95.0%;  Pred. No. 3.3e-35;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||     ||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101


RESULT 9
ABO44603
ID   ABO44603 standard; protein; 103 AA.
XX
AC   ABO44603;
XX
DT   02-OCT-2003  (first entry)
XX
DE   Novel human secreted protein #46.
XX
KW   Human; gene therapy; autoimmune disorder; multiple sclerosis; cancer;
KW   systemic lupus erythematosus; haematopoietic cell disorder; allergy;
KW   agammaglobulinaemia; ataxia telangiectasia; blood coagulation disorder;
KW   afibrinogenaemia; thrombocytopenia; graft-versus-host disease; arthritis;
KW   inflammatory condition; ischaemia-reperfusion injury; infectious disease;
KW   hyperproliferative disorder; purpura; viral infection; regeneration;
KW   bacterial infection; ulcer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   US2003065160-A1.
XX
PD   03-APR-2003.
XX
PF   07-DEC-2001; 2001US-00004860.
XX
```

PR 06-JUN-1997; 97US-0048875P.
PR 06-JUN-1997; 97US-0048876P.
PR 06-JUN-1997; 97US-0048877P.
PR 06-JUN-1997; 97US-0048878P.
PR 06-JUN-1997; 97US-0048880P.
PR 06-JUN-1997; 97US-0048881P.
PR 06-JUN-1997; 97US-0048882P.
PR 06-JUN-1997; 97US-0048883P.
PR 06-JUN-1997; 97US-0048884P.
PR 06-JUN-1997; 97US-0048885P.
PR 06-JUN-1997; 97US-0048892P.
PR 06-JUN-1997; 97US-0048893P.
PR 06-JUN-1997; 97US-0048894P.
PR 06-JUN-1997; 97US-0048895P.
PR 06-JUN-1997; 97US-0048896P.
PR 06-JUN-1997; 97US-0048897P.
PR 06-JUN-1997; 97US-0048898P.
PR 06-JUN-1997; 97US-0048899P.
PR 06-JUN-1997; 97US-0048900P.
PR 06-JUN-1997; 97US-0048901P.
PR 06-JUN-1997; 97US-0048915P.
PR 06-JUN-1997; 97US-0048916P.
PR 06-JUN-1997; 97US-0048917P.
PR 06-JUN-1997; 97US-0048949P.
PR 06-JUN-1997; 97US-0048962P.
PR 06-JUN-1997; 97US-0048963P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048970P.
PR 06-JUN-1997; 97US-0048971P.
PR 06-JUN-1997; 97US-0048972P.
PR 06-JUN-1997; 97US-0048974P.
PR 06-JUN-1997; 97US-0049019P.
PR 06-JUN-1997; 97US-0049020P.
PR 06-JUN-1997; 97US-0049373P.
PR 06-JUN-1997; 97US-0049374P.
PR 06-JUN-1997; 97US-0049375P.
PR 05-SEP-1997; 97US-0057584P.
PR 05-SEP-1997; 97US-0057627P.
PR 05-SEP-1997; 97US-0057628P.
PR 05-SEP-1997; 97US-0057629P.
PR 05-SEP-1997; 97US-0057634P.
PR 05-SEP-1997; 97US-0057635P.
PR 05-SEP-1997; 97US-0057642P.
PR 05-SEP-1997; 97US-0057643P.
PR 05-SEP-1997; 97US-0057644P.
PR 05-SEP-1997; 97US-0057645P.
PR 05-SEP-1997; 97US-0057646P.
PR 05-SEP-1997; 97US-0057647P.
PR 05-SEP-1997; 97US-0057648P.
PR 05-SEP-1997; 97US-0057649P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057651P.
PR 05-SEP-1997; 97US-0057654P.
PR 05-SEP-1997; 97US-0057661P.
PR 05-SEP-1997; 97US-0057662P.
PR 05-SEP-1997; 97US-0057666P.
PR 05-SEP-1997; 97US-0057667P.
PR 05-SEP-1997; 97US-0057668P.
PR 05-SEP-1997; 97US-0057760P.
PR 05-SEP-1997; 97US-0057761P.
PR 05-SEP-1997; 97US-0057762P.
PR 05-SEP-1997; 97US-0057763P.
PR 05-SEP-1997; 97US-0057764P.
PR 05-SEP-1997; 97US-0057765P.
PR 05-SEP-1997; 97US-0057769P.
PR 05-SEP-1997; 97US-0057770P.
PR 05-SEP-1997; 97US-0057771P.
PR 05-SEP-1997; 97US-0057774P.
PR 05-SEP-1997; 97US-0057775P.
PR 05-SEP-1997; 97US-0057776P.
PR 05-SEP-1997; 97US-0057777P.
PR 05-SEP-1997; 97US-0057778P.
PR 18-DEC-1997; 97US-0070923P.
PR 04-JUN-1998; 98WO-US011422.
PR 15-JUL-1998; 98US-0092921P.
PR 30-JUL-1998; 98US-0094657P.
PR 04-DEC-1998; 98US-00205258.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;
PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;
PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;
PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;
PI Carter KC;
XX
DR WPI; 2003-540804/51.
DR N-PSDB; ACH04740.
XX
PT New isolated protein, useful for preparing a composition for diagnosing
PT or treating cancer, inflammatory, immune or infectious diseases.
XX
PS Disclosure; SEQ ID NO 294; 172pp; English.
XX
CC The invention relates to an isolated HEMAE80 protein. The protein is
CC useful for preparing a composition for diagnosing or treating autoimmune
CC disorders e.g. multiple sclerosis and systemic lupus erythematosus;
CC haematopoietic cell disorders e.g. agammaglobulinaemia and ataxia
CC telangiectasia; blood coagulation disorders e.g. afibrinogenaemia and
CC thrombocytopenia; allergy; graft-versus-host disease; inflammatory
CC conditions e.g. ischaemia-reperfusion injury and arthritis;
CC hyperproliferative disorders e.g. cancer and purpura; infectious disease
CC e.g. viral infection and bacterial infection. The polynucleotide or
CC protein can be used to regenerate damaged tissue e.g. ulcers and
CC Alzheimer's disease. The present sequence represents the amino acid
CC sequence of a novel human secreted protein. Note: The sequence data for
CC this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030065160
XX
SQ Sequence 103 AA;

Query Match 33.5%; Score 477.5; DB 6; Length 103;
Best Local Similarity 95.0%; Pred. No. 3.3e-35;
Matches 96; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVPFPIMGLPGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||  |||||||||||||||||||||
Db         61 YALVPVPFPIMGLPGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101
```

RESULT 10
ABO26083
ID ABO26083 standard; protein; 103 AA.
XX
AC ABO26083;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human protein from novel secreted protein gene 46.
XX
KW Human; secreted protein; precerebellin-like protein;
KW neurodegenerative disorder; behavioural disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia; mania;
KW dementia; paranoia; psychosis; autism; immune disorder; infection;
KW inflammation; allergy; liver disorder; hepatoblastoma; jaundice;
KW hepatitis; immunological disorder; AIDS; leukaemia; rheumatoid arthritis;
KW sepsis; acne; psoriasis; cancer.
XX
OS Homo sapiens.
XX
PN US6525174-B1.

XX
PD 25-FEB-2003.
XX
PF 04-DEC-1998; 98US-00205258.
XX
PR 06-JUN-1997; 97US-0048875P.
PR 06-JUN-1997; 97US-0048876P.
PR 06-JUN-1997; 97US-0048877P.
PR 06-JUN-1997; 97US-0048878P.
PR 06-JUN-1997; 97US-0048880P.
PR 06-JUN-1997; 97US-0048881P.
PR 06-JUN-1997; 97US-0048882P.
PR 06-JUN-1997; 97US-0048883P.
PR 06-JUN-1997; 97US-0048884P.
PR 06-JUN-1997; 97US-0048885P.
PR 06-JUN-1997; 97US-0048892P.
PR 06-JUN-1997; 97US-0048893P.
PR 06-JUN-1997; 97US-0048894P.
PR 06-JUN-1997; 97US-0048895P.
PR 06-JUN-1997; 97US-0048896P.
PR 06-JUN-1997; 97US-0048897P.
PR 06-JUN-1997; 97US-0048898P.
PR 06-JUN-1997; 97US-0048899P.
PR 06-JUN-1997; 97US-0048900P.
PR 06-JUN-1997; 97US-0048901P.
PR 06-JUN-1997; 97US-0048915P.
PR 06-JUN-1997; 97US-0048916P.
PR 06-JUN-1997; 97US-0048917P.
PR 06-JUN-1997; 97US-0048949P.
PR 06-JUN-1997; 97US-0048962P.
PR 06-JUN-1997; 97US-0048963P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048970P.
PR 06-JUN-1997; 97US-0048971P.
PR 06-JUN-1997; 97US-0048972P.
PR 06-JUN-1997; 97US-0048974P.
PR 06-JUN-1997; 97US-0049019P.
PR 06-JUN-1997; 97US-0049020P.
PR 06-JUN-1997; 97US-0049373P.
PR 06-JUN-1997; 97US-0049374P.
PR 06-JUN-1997; 97US-0049375P.
PR 05-SEP-1997; 97US-0057584P.
PR 05-SEP-1997; 97US-0057627P.
PR 05-SEP-1997; 97US-0057628P.
PR 05-SEP-1997; 97US-0057629P.
PR 05-SEP-1997; 97US-0057634P.
PR 05-SEP-1997; 97US-0057635P.
PR 05-SEP-1997; 97US-0057642P.
PR 05-SEP-1997; 97US-0057643P.
PR 05-SEP-1997; 97US-0057644P.
PR 05-SEP-1997; 97US-0057645P.
PR 05-SEP-1997; 97US-0057646P.
PR 05-SEP-1997; 97US-0057647P.
PR 05-SEP-1997; 97US-0057648P.
PR 05-SEP-1997; 97US-0057649P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057651P.
PR 05-SEP-1997; 97US-0057654P.
PR 05-SEP-1997; 97US-0057661P.
PR 05-SEP-1997; 97US-0057662P.
PR 05-SEP-1997; 97US-0057666P.
PR 05-SEP-1997; 97US-0057667P.
PR 05-SEP-1997; 97US-0057668P.
PR 05-SEP-1997; 97US-0057760P.
PR 05-SEP-1997; 97US-0057761P.
PR 05-SEP-1997; 97US-0057762P.
PR 05-SEP-1997; 97US-0057763P.
PR 05-SEP-1997; 97US-0057764P.
PR 05-SEP-1997; 97US-0057765P.
PR 05-SEP-1997; 97US-0057769P.
PR 05-SEP-1997; 97US-0057770P.
PR 05-SEP-1997; 97US-0057771P.
PR 05-SEP-1997; 97US-0057774P.
PR 05-SEP-1997; 97US-0057775P.
PR 05-SEP-1997; 97US-0057776P.
PR 05-SEP-1997; 97US-0057777P.
PR 05-SEP-1997; 97US-0057778P.
PR 18-DEC-1997; 97US-0070923P.
PR 04-JUN-1998; 98WO-US011422.
PR 15-JUL-1998; 98US-0092921P.
PR 30-JUL-1998; 98US-0094657P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;
PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;
PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;
PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;
PI Carter KC;
XX
DR WPI; 2003-511926/48.
DR N-PSDB; ACD44550.
XX
PT New precerebellin-like protein, useful for diagnosing or treating
PT neurodegenerative and behavioral disorders, immune disorders, liver
PT disorders, and cancer.
XX
PS Disclosure; SEQ ID NO 294; 156pp; English.
XX
CC The invention relates to an isolated protein comprising amino acid
CC residues 33-205 or 1-205 of a a novel human secreted protein appearing as
CC ABO26252. The protein is encoded by one of 238 disclosed cDNA sequences
CC encoding 238 secreted proteins. ABO26252 is a precerebellin-like protein.
CC Also included are a composition comprising the protein and a carrier and
CC an isolated protein produced by expressing the protein cited above by a
CC cell, and recovering the protein. The proteins are useful for diagnosing
CC or treating neurodegenerative and behavioural disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, schizophrenia, mania,
CC dementia, paranoia, psychoses or autism), immune disorders (e.g.
CC infection, inflammation, allergy), liver disorders (e.g. hepatoblastoma,
CC jaundice, hepatitis), immunological disorders (e.g. AIDS, leukaemia,
CC rheumatoid arthritis, sepsis, acne, psoriasis) and cancer. The present
CC sequence is one of the 238 disclosed novel secreted proteins. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at:- seqdata.uspto.gov/sequence.html?DocID=6525174B1
XX
SQ Sequence 103 AA;

Query Match 33.5%; Score 477.5; DB 7; Length 103;
Best Local Similarity 95.0%; Pred. No. 3.3e-35;
Matches 96; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||     ||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101
```

RESULT 11
ABB50685
ID ABB50685 standard; protein; 181 AA.
XX
AC ABB50685;
XX
DT 07-FEB-2002 (first entry)
XX
DE Human secreted protein encoded by gene 46 SEQ ID NO:633.
XX
KW Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
KW dermatological; immunosuppressive; antiinflammatory; immunostimulant;
KW cytostatic; cardiant; vascular; anti-angiogenic; ophthalmological;
KW neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnerary;

KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis;
KW Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder;
KW corneal graft neovascularisation; diabetic retinopathy; regeneration;
KW neurological disorder; Huntington's chorea; Alzheimer's disease;
KW Parkinson's disease; infectious disease.
XX
OS Homo sapiens.
XX
PN WO200162891-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US005614.
XX
PR 24-FEB-2000; 2000US-0184836P.
PR 29-MAR-2000; 2000US-0193170P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Ebner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;
PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrie AM, Fan P;
PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G, Zeng Z;
PI Greene JM;
XX
DR WPI; 2001-625724/72.
XX
PT Nucleic acids encoding 207 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease
PT and diabetic retinopathy.
XX
PS Disclosure; Page 107; 1533pp; English.
XX
CC ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
CC activities based on the tissues and cells the genes are expressed in.
CC Example of these activities include: immunomodulatory; antisclerotic;
CC dermatological; immunosuppressive; antiinflammatory; immunostimulant;
CC anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;
CC neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;
CC antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used
CC in gene therapy and vaccine production. (I) and (II) can be used in the
CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC disorders (e.g. corneal graft neovascularisation and diabetic
CC retinopathy), neurological disorders (e.g. Huntington's chorea,
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
CC ABA83193 and ABB50300 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 181 AA;

```
  Query Match             23.3%;  Score 332;  DB 4;  Length 181;
  Best Local Similarity   85.5%;  Pred. No. 1.1e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
              :::||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy        157 SPPVSPGPLSPGGTPG 172
              |||||||    ||  |
Db        136 SPPVSPGLCHQGGRQG 151
```

RESULT 12

AAE20965
ID AAE20965 standard; protein; 181 AA.
XX
AC AAE20965;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human TR21 receptor protein fragment.
XX
KW Human; tumour necrosis factor receptor; TNF; TR21; TR22; immune disorder;
KW autoimmune haemolytic anaemia; rheumatoid arthritis; Addison's disease;
KW allergy; cancer; ulcerative colitis; cardiovascular disorder; epilepsy;
KW myocardial ischaemia; wound healing; neurological disease; infection;
KW cerebral anoxia; gene therapy; immunosuppressive; antiproliferative;
KW cytostatic; cardiant; vasotropic; neuroprotective; antibacterial;
KW virucide; fungicide; nootropic; ophthalmological.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 56
FT /label= Unknown
FT Misc-difference 57
FT /label= Unknown
FT Misc-difference 60
FT /label= Unknown
FT Misc-difference 83
FT /label= Unknown
FT Misc-difference 84
FT /label= Unknown
FT Misc-difference 165
FT /label= Unknown
XX
PN WO200207762-A1.
XX
PD 31-JAN-2002.
XX
PF 23-JUL-2001; 2001WO-US023124.
XX
PR 24-JUL-2000; 2000US-0220116P.
PR 27-JUL-2000; 2000US-0221143P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Zeng Z, Ruben SM, Rosen CA;
XX
DR WPI; 2002-171948/22.
XX
PT Two novel proteins, TR21 and TR22, which are members of the tumor
PT necrosis factor receptor, useful for the diagnosis and treatment of
PT immune disorders, cancer, cardiovascular disorders.
XX
PS Disclosure; Page 5; 248pp; English.
XX
CC The present invention relates to novel human tumour necrosis factor (TNF)
CC receptors, TR21 and TR22 and polynucleotides encoding them. Sequences of
CC the invention are useful in the diagnosis, treatment and prevention of
CC cancers (e.g., cancers of the adrenal gland, bone, urogenital or bone
CC marrow, in particular breast and ovarian cancer), immune disorders (e.g.,
CC autoimmune haemolytic anaemia, rheumatoid arthritis, allergies, Addison's
CC disease, ulcerative colitis), cardiovascular disorders (e.g., myocardial
CC ischaemias), wound healing, neurological diseases (e.g., cerebral anoxia,
CC epilepsy) and infectious diseases such as viral, bacterial, fungal and
CC parasitic infections. They are also useful in gene therapy. The present
CC sequence is human TR21 protein fragment
XX
SQ Sequence 181 AA;

```
  Query Match             23.3%;  Score 332;  DB 5;  Length 181;
  Best Local Similarity   85.5%;  Pred. No. 1.1e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
```

```
            :::||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy        157 SPPVSPGPLSPGGTPG 172
              |||||||    ||  |
Db        136 SPPVSPGLCHQGGRQG 151


RESULT 13
ABO44942
ID   ABO44942 standard; protein; 181 AA.
XX
AC   ABO44942;
XX
DT   02-OCT-2003  (first entry)
XX
DE   Novel human secreted protein #46 fragment #2.
XX
KW   Human; gene therapy; autoimmune disorder; multiple sclerosis; cancer;
KW   systemic lupus erythematosus; haematopoietic cell disorder; allergy;
KW   agammaglobulinaemia; ataxia telangiectasia; blood coagulation disorder;
KW   afibrinogenaemia; thrombocytopenia; graft-versus-host disease; arthritis;
KW   inflammatory condition; ischaemia-reperfusion injury; infectious disease;
KW   hyperproliferative disorder; purpura; viral infection; regeneration;
KW   bacterial infection; ulcer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   US2003065160-A1.
XX
PD   03-APR-2003.
XX
PF   07-DEC-2001; 2001US-00004860.
XX
PR   06-JUN-1997;   97US-0048875P.
PR   06-JUN-1997;   97US-0048876P.
PR   06-JUN-1997;   97US-0048877P.
PR   06-JUN-1997;   97US-0048878P.
PR   06-JUN-1997;   97US-0048880P.
PR   06-JUN-1997;   97US-0048881P.
PR   06-JUN-1997;   97US-0048882P.
PR   06-JUN-1997;   97US-0048883P.
PR   06-JUN-1997;   97US-0048884P.
PR   06-JUN-1997;   97US-0048885P.
PR   06-JUN-1997;   97US-0048892P.
PR   06-JUN-1997;   97US-0048893P.
PR   06-JUN-1997;   97US-0048894P.
PR   06-JUN-1997;   97US-0048895P.
PR   06-JUN-1997;   97US-0048896P.
PR   06-JUN-1997;   97US-0048897P.
PR   06-JUN-1997;   97US-0048898P.
PR   06-JUN-1997;   97US-0048899P.
PR   06-JUN-1997;   97US-0048900P.
PR   06-JUN-1997;   97US-0048901P.
PR   06-JUN-1997;   97US-0048915P.
PR   06-JUN-1997;   97US-0048916P.
PR   06-JUN-1997;   97US-0048917P.
PR   06-JUN-1997;   97US-0048949P.
PR   06-JUN-1997;   97US-0048962P.
PR   06-JUN-1997;   97US-0048963P.
PR   06-JUN-1997;   97US-0048964P.
PR   06-JUN-1997;   97US-0048970P.
PR   06-JUN-1997;   97US-0048971P.
PR   06-JUN-1997;   97US-0048972P.
PR   06-JUN-1997;   97US-0048974P.
PR   06-JUN-1997;   97US-0049019P.
PR   06-JUN-1997;   97US-0049020P.
PR   06-JUN-1997;   97US-0049373P.
PR   06-JUN-1997;   97US-0049374P.
PR   06-JUN-1997;   97US-0049375P.
PR   05-SEP-1997;   97US-0057584P.
PR   05-SEP-1997;   97US-0057627P.
PR   05-SEP-1997;   97US-0057628P.
PR   05-SEP-1997;   97US-0057629P.
PR   05-SEP-1997;   97US-0057634P.
PR   05-SEP-1997;   97US-0057635P.
PR   05-SEP-1997;   97US-0057642P.
PR   05-SEP-1997;   97US-0057643P.
PR   05-SEP-1997;   97US-0057644P.
PR   05-SEP-1997;   97US-0057645P.
PR   05-SEP-1997;   97US-0057646P.
PR   05-SEP-1997;   97US-0057647P.
PR   05-SEP-1997;   97US-0057648P.
PR   05-SEP-1997;   97US-0057649P.
PR   05-SEP-1997;   97US-0057650P.
PR   05-SEP-1997;   97US-0057651P.
PR   05-SEP-1997;   97US-0057654P.
PR   05-SEP-1997;   97US-0057661P.
PR   05-SEP-1997;   97US-0057662P.
PR   05-SEP-1997;   97US-0057666P.
PR   05-SEP-1997;   97US-0057667P.
PR   05-SEP-1997;   97US-0057668P.
PR   05-SEP-1997;   97US-0057760P.
PR   05-SEP-1997;   97US-0057761P.
PR   05-SEP-1997;   97US-0057762P.
PR   05-SEP-1997;   97US-0057763P.
PR   05-SEP-1997;   97US-0057764P.
PR   05-SEP-1997;   97US-0057765P.
PR   05-SEP-1997;   97US-0057769P.
PR   05-SEP-1997;   97US-0057770P.
PR   05-SEP-1997;   97US-0057771P.
PR   05-SEP-1997;   97US-0057774P.
PR   05-SEP-1997;   97US-0057775P.
PR   05-SEP-1997;   97US-0057776P.
PR   05-SEP-1997;   97US-0057777P.
PR   05-SEP-1997;   97US-0057778P.
PR   18-DEC-1997;   97US-0070923P.
PR   04-JUN-1998;   98WO-US011422.
PR   15-JUL-1998;   98US-0092921P.
PR   30-JUL-1998;   98US-0094657P.
PR   04-DEC-1998;   98US-00205258.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Young P,  Greene JM,  Ferrie AM,  Ruben SM,  Rosen CA,  Hu J;
PI   Olsen HS,  Ebner R,  Brewer LA,  Moore PA,  Shi Y,  Florence C;
PI   Florence K,  Lafleur DW,  Ni J,  Fan P,  Wei Y,  Fischer CL,  Soppet DR;
PI   Li Y,  Zeng Z,  Kyaw H,  Yu G,  Feng P,  Dillon PJ,  Endress GA;
PI   Carter KC;
XX
DR   WPI; 2003-540804/51.
XX
PT   New isolated protein, useful for preparing a composition for diagnosing
PT   or treating cancer, inflammatory, immune or infectious diseases.
XX
PS   Disclosure; Page 28; 172pp; English.
XX
CC   The invention relates to an isolated HEMAE80 protein. The protein is
CC   useful for preparing a composition for diagnosing or treating autoimmune
CC   disorders e.g. multiple sclerosis and systemic lupus erythematosus;
CC   haematopoietic cell disorders e.g. agammaglobulinaemia and ataxia
CC   telangiectasia; blood coagulation disorders e.g. afibrinogenaemia and
CC   thrombocytopenia; allergy; graft-versus-host disease; inflammatory
CC   conditions e.g. ischaemia-reperfusion injury and arthritis;
CC   hyperproliferative disorders e.g. cancer and purpura; infectious disease
CC   e.g. viral infection and bacterial infection. The polynucleotide or
CC   protein can be used to regenerate damaged tissue e.g. ulcers and
CC   Alzheimer's disease. The present sequence represents the amino acid
CC   sequence of a novel human secreted protein fragment. Note: The sequence
CC   data for this patent did not form part of the printed specification but
CC   was obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030065160
XX
SQ   Sequence 181 AA;

  Query Match             23.3%;  Score 332;  DB 6;  Length 181;
```

```
  Best Local Similarity   85.5%;  Pred. No. 1.1e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
               :::|||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy         157 SPPVSPGPLSPGGTPG 172
               |||||||    ||  |
Db         136 SPPVSPGLCHQGGRQG 151


RESULT 14
ABO26422
ID   ABO26422 standard; protein; 181 AA.
XX
AC   ABO26422;
XX
DT   10-SEP-2003  (first entry)
XX
DE   Protein associated with novel secreted protein gene 46 #2.
XX
KW   Secreted protein; precerebellin-like protein; sepsis; acne; psoriasis;
KW   neurodegenerative disorder; behavioural disorder; Alzheimer's disease;
KW   Parkinson's disease; Huntington's disease; schizophrenia; mania;
KW   dementia; paranoia; psychosis; autism; immune disorder; infection;
KW   inflammation; allergy; liver disorder; hepatoblastoma; jaundice;
KW   hepatitis; immunological disorder; AIDS; leukaemia; rheumatoid arthritis;
KW   cancer.
XX
OS   Unidentified.
XX
PN   US6525174-B1.
XX
PD   25-FEB-2003.
XX
PF   04-DEC-1998;   98US-00205258.
XX
PR   06-JUN-1997;   97US-0048875P.
PR   06-JUN-1997;   97US-0048876P.
PR   06-JUN-1997;   97US-0048877P.
PR   06-JUN-1997;   97US-0048878P.
PR   06-JUN-1997;   97US-0048880P.
PR   06-JUN-1997;   97US-0048881P.
PR   06-JUN-1997;   97US-0048882P.
PR   06-JUN-1997;   97US-0048883P.
PR   06-JUN-1997;   97US-0048884P.
PR   06-JUN-1997;   97US-0048885P.
PR   06-JUN-1997;   97US-0048892P.
PR   06-JUN-1997;   97US-0048893P.
PR   06-JUN-1997;   97US-0048894P.
PR   06-JUN-1997;   97US-0048895P.
PR   06-JUN-1997;   97US-0048896P.
PR   06-JUN-1997;   97US-0048897P.
PR   06-JUN-1997;   97US-0048898P.
PR   06-JUN-1997;   97US-0048899P.
PR   06-JUN-1997;   97US-0048900P.
PR   06-JUN-1997;   97US-0048901P.
PR   06-JUN-1997;   97US-0048915P.
PR   06-JUN-1997;   97US-0048916P.
PR   06-JUN-1997;   97US-0048917P.
PR   06-JUN-1997;   97US-0048949P.
PR   06-JUN-1997;   97US-0048962P.
PR   06-JUN-1997;   97US-0048963P.
PR   06-JUN-1997;   97US-0048964P.
PR   06-JUN-1997;   97US-0048970P.
PR   06-JUN-1997;   97US-0048971P.
PR   06-JUN-1997;   97US-0048972P.
PR   06-JUN-1997;   97US-0048974P.
PR   06-JUN-1997;   97US-0049019P.
PR   06-JUN-1997;   97US-0049020P.
PR   06-JUN-1997;   97US-0049373P.
PR   06-JUN-1997;   97US-0049374P.
PR   06-JUN-1997;   97US-0049375P.
PR   05-SEP-1997;   97US-0057584P.
PR   05-SEP-1997;   97US-0057627P.
PR   05-SEP-1997;   97US-0057628P.
PR   05-SEP-1997;   97US-0057629P.
PR   05-SEP-1997;   97US-0057634P.
PR   05-SEP-1997;   97US-0057635P.
PR   05-SEP-1997;   97US-0057642P.
PR   05-SEP-1997;   97US-0057643P.
PR   05-SEP-1997;   97US-0057644P.
PR   05-SEP-1997;   97US-0057645P.
PR   05-SEP-1997;   97US-0057646P.
PR   05-SEP-1997;   97US-0057647P.
PR   05-SEP-1997;   97US-0057648P.
PR   05-SEP-1997;   97US-0057649P.
PR   05-SEP-1997;   97US-0057650P.
PR   05-SEP-1997;   97US-0057651P.
PR   05-SEP-1997;   97US-0057654P.
PR   05-SEP-1997;   97US-0057661P.
PR   05-SEP-1997;   97US-0057662P.
PR   05-SEP-1997;   97US-0057666P.
PR   05-SEP-1997;   97US-0057667P.
PR   05-SEP-1997;   97US-0057668P.
PR   05-SEP-1997;   97US-0057760P.
PR   05-SEP-1997;   97US-0057761P.
PR   05-SEP-1997;   97US-0057762P.
PR   05-SEP-1997;   97US-0057763P.
PR   05-SEP-1997;   97US-0057764P.
PR   05-SEP-1997;   97US-0057765P.
PR   05-SEP-1997;   97US-0057769P.
PR   05-SEP-1997;   97US-0057770P.
PR   05-SEP-1997;   97US-0057771P.
PR   05-SEP-1997;   97US-0057774P.
PR   05-SEP-1997;   97US-0057775P.
PR   05-SEP-1997;   97US-0057776P.
PR   05-SEP-1997;   97US-0057777P.
PR   05-SEP-1997;   97US-0057778P.
PR   18-DEC-1997;   97US-0070923P.
PR   04-JUN-1998;   98WO-US011422.
PR   15-JUL-1998;   98US-0092921P.
PR   30-JUL-1998;   98US-0094657P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Young P,  Greene JM,  Ferrie AM,  Ruben SM,  Rosen CA,  Hu J;
PI   Olsen HS,  Ebner R,  Brewer LA,  Moore PA,  Shi Y,  Florence C;
PI   Florence K,  Lafleur DW,  Ni J,  Fan P,  Wei Y,  Fischer CL,  Soppet DR;
PI   Li Y,  Zeng Z,  Kyaw H,  Yu G,  Feng P,  Dillon PJ,  Endress GA;
PI   Carter KC;
XX
DR   WPI; 2003-511926/48.
XX
PT   New precerebellin-like protein, useful for diagnosing or treating
PT   neurodegenerative and behavioral disorders, immune disorders, liver
PT   disorders, and cancer.
XX
PS   Disclosure; Col 52; 156pp; English.
XX
CC   The invention relates to an isolated protein comprising amino acid
CC   residues 33-205 or 1-205 of a a novel human secreted protein appearing as
CC   ABO26252. The protein is encoded by one of 238 disclosed cDNA sequences
CC   encoding 238 secreted proteins. ABO26252 is a precerebellin-like protein.
CC   Also included are a composition comprising the protein and a carrier and
CC   an isolated protein produced by expressing the protein cited above by a
CC   cell, and recovering the protein. The proteins are useful for diagnosing
CC   or treating neurodegenerative and behavioural disorders (e.g. Alzheimer's
CC   disease, Parkinson's disease, Huntington's disease, schizophrenia, mania,
CC   dementia, paranoia, psychoses or autism), immune disorders (e.g.
CC   infection, inflammation, allergy), liver disorders (e.g. hepatoblastoma,
CC   jaundice, hepatitis), immunological disorders (e.g. AIDS, leukaemia,
CC   rheumatoid arthritis, sepsis, acne, psoriasis) and cancer. The present
CC   sequence is a protein associated with one of the 238 disclosed novel
CC   secreted proteins
```

```
XX
SQ   Sequence 181 AA;

  Query Match             23.3%;  Score 332;  DB 7;  Length 181;
  Best Local Similarity   85.5%;  Pred. No. 1.1e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
               :::|||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy         157 SPPVSPGPLSPGGTPG 172
               |||||||    ||  |
Db         136 SPPVSPGLCHQGGRQG 151


RESULT 15
ADP29453
ID   ADP29453 standard; protein; 382 AA.
XX
AC   ADP29453;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Human secreted protein SEQ ID #220.
XX
KW   Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW   cancer; inflammatory; immune; human secreted protein.
XX
OS   Homo sapiens.
XX
PN   WO2004035732-A2.
XX
PD   29-APR-2004.
XX
PF   28-AUG-2003; 2003WO-US026780.
XX
PR   29-AUG-2002; 2002US-0406576P.
PR   29-AUG-2002; 2002US-0406579P.
PR   29-AUG-2002; 2002US-0406585P.
PR   29-AUG-2002; 2002US-0406588P.
PR   29-AUG-2002; 2002US-0406608P.
PR   29-AUG-2002; 2002US-0406611P.
PR   29-AUG-2002; 2002US-0406612P.
PR   29-AUG-2002; 2002US-0406616P.
PR   29-AUG-2002; 2002US-0406640P.
PR   29-AUG-2002; 2002US-0406642P.
PR   29-AUG-2002; 2002US-0406646P.
PR   29-AUG-2002; 2002US-0406653P.
PR   29-AUG-2002; 2002US-0406655P.
PR   29-AUG-2002; 2002US-0406666P.
PR   17-SEP-2002; 2002US-0410946P.
PR   17-SEP-2002; 2002US-0410947P.
PR   17-SEP-2002; 2002US-0410948P.
PR   17-SEP-2002; 2002US-0410949P.
PR   17-SEP-2002; 2002US-0410953P.
PR   17-SEP-2002; 2002US-0410957P.
PR   17-SEP-2002; 2002US-0410958P.
PR   17-SEP-2002; 2002US-0410959P.
PR   17-SEP-2002; 2002US-0410960P.
PR   17-SEP-2002; 2002US-0410961P.
PR   17-SEP-2002; 2002US-0410962P.
PR   17-SEP-2002; 2002US-0411019P.
PR   17-SEP-2002; 2002US-0411022P.
PR   17-SEP-2002; 2002US-0411023P.
PR   17-SEP-2002; 2002US-0411024P.
PR   17-SEP-2002; 2002US-0411032P.
PR   17-SEP-2002; 2002US-0411035P.
PR   17-SEP-2002; 2002US-0411037P.
PR   17-SEP-2002; 2002US-0411041P.
PR   17-SEP-2002; 2002US-0411045P.
PR   17-SEP-2002; 2002US-0411046P.
PR   17-SEP-2002; 2002US-0411048P.
PR   17-SEP-2002; 2002US-0411052P.
PR   17-SEP-2002; 2002US-0411055P.
PR   17-SEP-2002; 2002US-0411073P.
PR   17-SEP-2002; 2002US-0411082P.
PR   17-SEP-2002; 2002US-0411101P.
PR   17-SEP-2002; 2002US-0411111P.
PR   18-APR-2003; 2003US-0463700P.
PR   18-APR-2003; 2003US-0463708P.
PR   18-APR-2003; 2003US-0463716P.
PR   18-APR-2003; 2003US-0463732P.
PR   02-MAY-2003; 2003US-0467199P.
PR   02-MAY-2003; 2003US-0467201P.
PR   02-MAY-2003; 2003US-0467203P.
PR   02-MAY-2003; 2003US-0467230P.
PR   19-MAY-2003; 2003US-0471306P.
PR   19-MAY-2003; 2003US-0471336P.
PR   22-MAY-2003; 2003US-0472420P.
PR   22-MAY-2003; 2003US-0472430P.
PR   09-JUN-2003; 2003US-0476609P.
PR   09-JUN-2003; 2003US-0476641P.
PR   08-JUL-2003; 2003US-0485218P.
PR   08-JUL-2003; 2003US-0485223P.
PR   08-JUL-2003; 2003US-0485224P.
PR   08-JUL-2003; 2003US-0485325P.
PR   14-JUL-2003; 2003US-0486446P.
PR   14-JUL-2003; 2003US-0486480P.
PR   15-JUL-2003; 2003US-0486891P.
PR   15-JUL-2003; 2003US-0486960P.
PR   08-AUG-2003; 2003US-0493341P.
PR   08-AUG-2003; 2003US-0493370P.
PR   08-AUG-2003; 2003US-0493573P.
PR   08-AUG-2003; 2003US-0493577P.
XX
PA   (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI   Williams LT,  Chu K,  Lee E,  Hestir K,  Beaurang PA,  Behrens D;
PI   Halenbeck RF,  Huang MM,  Kothakota S,  Haishan L,  Linnemann T;
PI   Pierce K,  Wang Y,  Wong JGP,  Wu G,  Zhang H;
XX
DR   WPI; 2004-348438/32.
XX
PT   New nucleic acid molecule for diagnosing, preventing or treating diseases
PT   such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT   genetic, bacterial and viral diseases.
XX
PS   Claim 1; SEQ ID NO 1451; 428pp; English.
XX
CC   The present invention relates to an isolated nucleic acid molecule
CC   encoding a polypeptide which is believed to be cytostatic,
CC   antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC   composition and methods are useful for diagnosing, preventing and
CC   treating diseases such as proliferative (e.g. cancer), inflammatory,
CC   immune, metabolic, genetic, bacterial and viral diseases. The present
CC   sequence represents a human secreted protein. The present sequence is
CC   available on WIPOWEB and is not in the specification.
XX
SQ   Sequence 382 AA;

  Query Match             21.7%;  Score 310;  DB 8;  Length 382;
  Best Local Similarity   32.1%;  Pred. No. 2.8e-19;
  Matches   84;  Conservative   36;  Mismatches   72;  Indels   70;  Gaps   12;

Qy          44 TPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
               |  | | |  | ||      |   |:|||| :||| |:|:|:|:||:||| ||   |
Db         102 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 161

Qy          94 ----QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESP 149
                   |:               | ||:| :|  | : : ||  |: :: :   |  :
Db         162 GGGGSGTED------------ANEDTIGVLVRLITEKKENAAALEELLKE---YHSKQL 205

Qy         150 VTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCRHKRWHPI-- 203
               |  |     |||  | :|  |  |:|  ||||||  |:  :    | ||  |:|  :
Db         206 VQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCSQKKWPEVLL 260
```

```
Qy        204 ----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLM 246
                              ||   :  ::  ||||:|:||||||| ::       |:|:
Db        261 SPEAVAATTPVPSLLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI------PEQRTSS 314

Qy        247 SVSGAETVN------GEVPATP 262
              ||  :|:       |::| :|
Db        315 MVSEVKTITEAGPSWGDLPDSP 336
```

Search completed: February 23, 2006, 12:03:53
Job time : 190 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 23, 2006, 12:09:38 ; Search time 18 Seconds
(without alignments)
224.136 Million cell updates/sec

Title: US-10-620-562-2
Perfect score: 1426
Sequence: 1 MAPRALPGSAVLAAAVFVGG.....ETVNGEVPATPVKRERSGTE 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 117670 seqs, 14887254 residues

Total number of hits satisfying chosen parameters: 117670

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	307	21.5	430	7	US-11-042-814-2	Sequence 2, Appli
2	295.5	20.7	436	7	US-11-042-814-4	Sequence 4, Appli
3	97.5	6.8	447	7	US-11-072-512-2688	Sequence 2688, Ap
4	97	6.8	287	7	US-11-198-819-6	Sequence 6, Appli
5	97	6.8	287	7	US-11-198-819-8	Sequence 8, Appli
6	87.5	6.1	948	6	US-10-485-517-131	Sequence 131, App
7	85.5	6.0	367	6	US-10-888-962-6	Sequence 6, Appli
8	85	6.0	241	7	US-11-072-512-2203	Sequence 2203, Ap
9	85	6.0	377	7	US-11-240-769-67	Sequence 67, Appl
10	84.5	5.9	954	6	US-10-467-962B-31	Sequence 31, Appl
11	84	5.9	1344	7	US-11-072-512-2452	Sequence 2452, Ap
12	82.5	5.8	399	6	US-10-467-657-1672	Sequence 1672, Ap
13	82.5	5.8	1140	7	US-11-169-232-104	Sequence 104, App
14	82.5	5.8	1363	7	US-11-169-232-52	Sequence 52, Appl
15	82.5	5.8	1404	7	US-11-169-232-2	Sequence 2, Appli
16	82.5	5.8	1404	7	US-11-169-232-62	Sequence 62, Appl
17	82	5.8	948	6	US-10-523-477-14	Sequence 14, Appl
18	81.5	5.7	626	7	US-11-072-512-2827	Sequence 2827, Ap
19	81	5.7	220	7	US-11-169-232-96	Sequence 96, Appl
20	80.5	5.6	714	7	US-11-124-367A-268	Sequence 268, App
21	79.5	5.6	254	6	US-10-714-887-126	Sequence 126, App
22	79.5	5.6	427	6	US-10-714-887-122	Sequence 122, App
23	79.5	5.6	619	7	US-11-052-554A-229	Sequence 229, App
24	79.5	5.6	2910	7	US-11-087-100-2	Sequence 2, Appli
25	79.5	5.6	2910	7	US-11-087-084-2	Sequence 2, Appli
26	79.5	5.6	2910	7	US-11-087-085-2	Sequence 2, Appli
27	78.5	5.5	294	6	US-10-485-788A-522	Sequence 522, App
28	78.5	5.5	831	6	US-10-467-657-4486	Sequence 4486, Ap
29	78	5.5	619	6	US-10-485-517-374	Sequence 374, App
30	78	5.5	2161	7	US-11-126-313-31	Sequence 31, Appl
31	78	5.5	2760	7	US-11-124-367A-444	Sequence 444, App
32	78	5.5	2803	7	US-11-124-367A-442	Sequence 442, App
33	78	5.5	2803	7	US-11-124-367A-445	Sequence 445, App
34	78	5.5	2984	7	US-11-124-367A-443	Sequence 443, App
35	78	5.5	3027	7	US-11-124-367A-441	Sequence 441, App
36	77	5.4	1972	7	US-11-124-367A-446	Sequence 446, App
37	76.5	5.4	431	7	US-11-058-735-79	Sequence 79, Appl
38	76.5	5.4	441	7	US-11-175-153-5	Sequence 5, Appli
39	76	5.3	227	7	US-11-176-830-762	Sequence 762, App
40	75.5	5.3	296	7	US-11-169-232-70	Sequence 70, Appl
41	75.5	5.3	828	6	US-10-995-561-983	Sequence 983, App
42	75.5	5.3	918	6	US-10-995-561-981	Sequence 981, App
43	75.5	5.3	1019	6	US-10-995-561-982	Sequence 982, App
44	75.5	5.3	1463	6	US-10-971-982-3	Sequence 3, Appli
45	75	5.3	274	6	US-10-353-783-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-11-042-814-2
; Sequence 2, Application US/11042814
; Publication No. US20060024267A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; APPLICANT: Welcher, Andrew A
; APPLICANT: Boedigheimer, Michael J
; APPLICANT: Shu, Junyan
; APPLICANT: Gary M. Fox
; TITLE OF INVENTION: TNFr/OPG-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017/36854
; CURRENT APPLICATION NUMBER: US/11/042,814
; CURRENT FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: US/10/146,574
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/724,037
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-042-814-2

Query Match 21.5%; Score 307; DB 7; Length 430;
Best Local Similarity 31.0%; Pred. No. 5.7e-21;
Matches 84; Conservative 36; Mismatches 69; Indels 82; Gaps 12;

```
Qy          44 TPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
               |  | | |  | ||      |   |:|||| :||| |:|:|:|||:||| ||   |
Db         144 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 203

Qy          94 ------------QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
                         :|                 | ||:| :|  | : : || |: :: :
Db         204 GGGGSGINPAYRTED---------------ANEDTIGVLVRLITEKKENAAALEELLKE 247

Qy         141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCR 196
                 |  :  |  |      |||  | :|   |  |:|   ||||||  |:  :    | ||
Db         248 ---YHSKQLVQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCS 299

Qy         197 HKRWHFI-----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKS 237
                 |:|  :                 ||   :  ::  ||||:|:|||||||| ::
Db         300 QKKWPEVLLSPEAVAATTPVPSLLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI---- 355

Qy         238 NQKERRSLMSVSGAETVN------GEVPATP 262
                 |:|:   ||  :|:       |::| :|
```

```
Db        356 --PEQRTSSMVSEVKTITEAGPSWGDLPDSP 384


RESULT 2
US-11-042-814-4
; Sequence 4, Application US/11042814
; Publication No. US20060024267A1
; GENERAL INFORMATION:
;  APPLICANT: Jing, Shuqian
;  APPLICANT:  Welcher, Andrew A
;  APPLICANT:  Boedigheimer, Michael J
;  APPLICANT:  Shu, Junyan
;  APPLICANT:  Gary M. Fox
;  TITLE OF INVENTION: TNFr/OPG-LIKE MOLECULES AND USES THEREOF
;  FILE REFERENCE: 01017/36854
;  CURRENT APPLICATION NUMBER: US/11/042,814
;  CURRENT FILING DATE:  2005-01-25
;  PRIOR APPLICATION NUMBER: US/10/146,574
;  PRIOR FILING DATE: 2002-05-15
;  PRIOR APPLICATION NUMBER: US/09/724,037
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 436
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-11-042-814-4

  Query Match             20.7%;  Score 295.5;  DB 7;  Length 436;
  Best Local Similarity   30.6%;  Pred. No. 6.8e-20;
  Matches   82;  Conservative   36;  Mismatches   89;  Indels   61;  Gaps   12;

Qy         31 NGSSRTLHSRTETTPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKG 85
              ||  |           | | |  | ||      |   |:|||| :||| |:|:|:|||:|| 
Db        148 NGEPR----------QPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKG 197

Qy         86 YRCTTEAEQDIEE----EKVEKIELNDSVNENSDTVGQIVHYIMKNEANA----DVLKAM 137
              | || : |           :      :  ||  ||:| :|  | : : ||    ::||  
Db        198 YHCTAQKEVGPSPGGGGSGINPAYRTEDANE--DTIGVLVRLITEKKENAAALEELLKEY 255

Qy        138 VADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCH 193
               :   :     ||     | || :             |:|   |||||| |:  :    | 
Db        256 HSKQLVQTSHRPVPRLLPASPSI-----------PHICPHHHHLHTVQGLASLSGPCCS 303

Qy        194 RCRHKRWHFI----------KPT---NKSR----ESRPRRQGEVTVLSVGRFRVTKVE 234
              ||  | |  :            ||     ||     ::| ||||:|:|||||||| :: 
Db        304 RCSQK-WPEVLLSPEAAAATTPAPTLLPTASRAPKASAKPGRQGEITILSVGRFRVARIP 362

Qy        235 HKSNQKERRSLMSVSGAETVNGEVPATP 262
               :       : ::: |    |::| :| 
Db        363 EQRTSSLLSEVKTITEAGPSEGDLPDSP 390


RESULT 3
US-11-072-512-2688
; Sequence 2688, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2688
;   LENGTH: 447
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-2688

  Query Match             6.8%;  Score 97.5;  DB 7;  Length 447;
  Best Local Similarity   20.6%;  Pred. No. 0.18;
  Matches   72;  Conservative   37;  Mismatches   87;  Indels  153;  Gaps   16;

Qy         29 PDNGSSRTLH----------------------------SRTETTPSPSNDTGNGHPEYI 59
              ||| :|:|                            :| :::|  |  : :  | | 
Db         73 PDNTTSKTTDCLQTKGFSNSTEHKRGSVAQKVQEFNNCLNRGQSSPQRSYSSSHSSPAKI 132

Qy         60 AYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTV-G 118
                                           | | |    ||  :  ::  : |  | : || 
Db        133 -------------------------QRATQEPVAKIEGTQESQMVGSSSTREKASTVLS 166

Qy        119 QIVHYIMKNEA---------------------NADVLKAM-----VADNSLY------- 144
              |||  |   ::                    :||| |:       |   ||:        
Db        167 QIVASIQPPQSPPETPQSGPKACSVEELYAIPPDADVAKSTPKSTPVRPKSLFTSQPSGE 226

Qy        145 ---------------DPE-SPVTPS------TPGSPPVSPGPLSPGGTPGKHVCGHHLH 181
                            ||   |||||      :| | | :| |  |  |  :  |: | 
Db        227 AEAPQTTDSPTTKVQKDPSIKPVTPSPSKLVTSPQSEPPAPFP-PPRSTSSPYHAGNLLQ 285

Qy        182 TVGGVVERDVCHRCRH-KRWHFIKPTNKSR----------------ESRPRRQGEVTVL 223
                          ||   |   |||: :|                :::|:|   ::   
Db        286 -------------RHFTNW--TKPTSPTRSTEAESVLHSEGSRRAADAKPKRW--ISFK 327

Qy        224 SVGRFRVTKVE-HKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
              | | | |  |  |  ::|:  |: : |  ||   :|  ||:|    ||  
Db        328 SFFRRRKTDEEDDKEKEREKGKLVGLDG--TVIHMLPPPPVQRHHWFTE 374


RESULT 4
US-11-198-819-6
; Sequence 6, Application US/11198819
; Publication No. US20050287582A1
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/11/198,819
;      FILING DATE:  04-AUG-2005
;      CLASSIFICATION:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  287 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-11-198-819-6

  Query Match             6.8%;  Score 97;  DB 7;  Length 287;
  Best Local Similarity   22.7%;  Pred. No. 0.11;
  Matches   58;  Conservative   45;  Mismatches  105;  Indels   48;  Gaps   14;

Qy         34 SRTLHSRTETTPSPSNDTGNGHPEYIAYALVPV-----FFIMGLFGVLICHLLKKKGYRC 88
              ::|:|:: |  | ||      |        ::|:     |   |  ||    | ::   |
Db         16 AQTIHTQEEDLPRPSISAEPG-------TVIPLGSHVTFVCRGPVGVQTFRLERES--RS 66

Qy         89 TTEAEQDIEEEKVEKIELN---DSVNENSDTVGQIVHY-IMKNEANADVLKAMVADNSLY 144
              |    :|: :    : |     |||:| :    : ::|   |    :| |: :| : |
Db         67 TYNDTEDVSQASPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSEQSDYLELLVKETS-G 125

Qy        145 DPESPVT-PSTPGSPPVSP--------GPLSPGGTPGKHVCGHHLHTVGGV-VERDVCHR 194
              |:|| | | :   |   |          | | |      :   ||: : || |    |
Db        126 GPDSPDTEPGSSAGPTQRPSDNSHNEHAPASQG------LKAEHLYILIGVSVVFLFCLL 179

Qy        195 -----CRHKRWHFIK---PTNKSRESRPRRQGEVTV----LSVGRFRVTKVEHKSNQKER 242
                  | | | : ||   | :|  | :|::: :: |     :  :  |  :  |  : :
Db        180 LLVLFCLH-RQNQIKQGPPRSKDEEQKPQQRPDLAVDVLERTADKATVNGLPEKDRETDT 238

Qy        243 RSLMSVSGAETVNGEV 258
              :| : |  |    ::
Db        239 SALAAGSSQEVTYAQL 254


RESULT 5
US-11-198-819-8
; Sequence 8, Application US/11198819
; Publication No. US20050287582A1
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/11/198,819
;      FILING DATE:  04-AUG-2005
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  287 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-11-198-819-8

  Query Match             6.8%;  Score 97;  DB 7;  Length 287;
  Best Local Similarity   22.7%;  Pred. No. 0.11;
  Matches   58;  Conservative   45;  Mismatches  105;  Indels   48;  Gaps   14;

Qy         34 SRTLHSRTETTPSPSNDTGNGHPEYIAYALVPV-----FFIMGLFGVLICHLLKKKGYRC 88
              ::|:|:: |  | ||      |        ::|:     |   |  ||    | ::   |
Db         16 AQTIHTQEEDLPRPSISAEPG-------TVIPLGSHVTFVCRGPVGVQTFRLERES--RS 66

Qy         89 TTEAEQDIEEEKVEKIELN---DSVNENSDTVGQIVHY-IMKNEANADVLKAMVADNSLY 144
              |    :|: :    : |     |||:| :    : ::|   |    :| |: :| : |
Db         67 TYNDTEDVSQASPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSEQSDYLELLVKETS-G 125

Qy        145 DPESPVT-PSTPGSPPVSP--------GPLSPGGTPGKHVCGHHLHTVGGV-VERDVCHR 194
              |:|| | | :   |   |          | | |      :   ||: : || |    |
Db        126 GPDSPDTEPGSSAGPTQRPSDNSHNEHAPASQG------LKAEHLYILIGVSVVFLFCLL 179

Qy        195 -----CRHKRWHFIK---PTNKSRESRPRRQGEVTV----LSVGRFRVTKVEHKSNQKER 242
                  | | | : ||   | :|  | :|::: :: |     :  :  |  :  |  : :
Db        180 LLVLFCLH-RQNQIKQGPPRSKDEEQKPQQRPDLAVDVLERTADKATVNGLPEKDRETDT 238

Qy        243 RSLMSVSGAETVNGEV 258
              :| : |  |    ::
Db        239 SALAAGSSQEVTYAQL 254


RESULT 6
US-10-485-517-131
; Sequence 131, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
;  APPLICANT: University of Sheffield
;  APPLICANT:  Biosynexus Incorporated
;  APPLICANT:  Foster, Simon
;  APPLICANT:  Mond, James
;  TITLE OF INVENTION: Antigenic Polypeptides
;  FILE REFERENCE: P100629WO
;  CURRENT APPLICATION NUMBER: US/10/485,517
;  CURRENT FILING DATE:  2004-02-02
;  PRIOR APPLICATION NUMBER: GB 0118825.9
;  PRIOR FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: GB 0200349.9
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 424
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 131
```

```
;   LENGTH: 948
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-485-517-131

  Query Match             6.1%;  Score 87.5;  DB 6;  Length 948;
  Best Local Similarity   21.2%;  Pred. No. 4.4;
  Matches   55;  Conservative   30;  Mismatches   94;  Indels   81;  Gaps   12;

Qy         26 LVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEY--------IAYALVPVFFIMGLFGVLI 77
              :|  |  |   :| :      |   ||    |:|        | :  ||
Db        714 IVDIDFDSVPQIHGQNNGNQSFEEDTEKDKPKYEQGGNIIDIDFDSVP----------- 761

Qy         78 CHLLKKKGYRCTTE-AEQDIEEEKVE-KIELNDSVNENSDTVGQIV-HYIMKNEANADVL 134
               |:    |:   ||  |:|  ::|   :   ::||:   ||: |:  |   :     |
Db        762 -HI---HGFNKHTEIIEEDTNKDKPNYQFGGHNSVDFEEDTLPQVSGHNEGQQTIEEDTT 817

Qy        135 KAMVADNSLYDPESPVTPSTPGSP------------PVSPGPLSP--GGTPGKHVCGHH 179
              :|       | :| ||  |  |              | :| | :|     |||
Db        818 PPIV-------PPTPPTPEVPSEPETPTPPTPEVPSEPETPTPPTPEVPTEPGKP----- 865

Qy        180 LHTVGGVVERDVCHRCRHKRWHFIKPTNK--SRESRPRRQGEVT--VLSVGRFRVTKVEH 235
                                          | |  :   : |:|  ||:|   |: :     |
Db        866 ----------------------IPPAKEEPKKPSKPVEQGKVVTPVIEINEKVKAVVPT 902

Qy        236 KSNQKERRSLMSVSGAETVN 255
              |  | ::  |    | |: |
Db        903 KKAQSKKSELPETGGEESTN 922


RESULT 7
US-10-888-962-6
; Sequence 6, Application US/10888962
; Publication No. US20050266531A1
; GENERAL INFORMATION:
;  APPLICANT: Saint Louis University
;  APPLICANT:  Ray, Ranjit
;  APPLICANT:  Yie-Hwa, Chang
;  APPLICANT:  Ray, Ratna
;  APPLICANT:  Basu, Arnab
;  TITLE OF INVENTION: Compositions and Methods for Inhibiting Liver Stellate Cell
;  TITLE OF INVENTION:  Growth
;  FILE REFERENCE: SLU 03-013 PCT
;  CURRENT APPLICATION NUMBER: US/10/888,962
;  CURRENT FILING DATE:  2004-07-09
;  PRIOR APPLICATION NUMBER: 60/487,126
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
;   LENGTH: 367
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-888-962-6

  Query Match             6.0%;  Score 85.5;  DB 6;  Length 367;
  Best Local Similarity   25.6%;  Pred. No. 1.8;
  Matches   58;  Conservative   23;  Mismatches   93;  Indels   53;  Gaps   13;

Qy         25 PLVAPDNGSSRTLHSRTETTP--SPSNDTGNGHPEYIAYA-LVPV-------FFIMGLFG 74
              ||:|| |  :| :|:       :  |:  |   | |: | |||:       | : |
Db        150 PLLAPLN-DTRVVHAAKAALAAFNAQNNGSNFQLEEISRAQLVPLPPSTYVEFTVSGTDC 208

Qy         75 V-------LICHLLKKKGYR-CTTEAEQDIEEEKVEKIELNDSVNENSDTVGQIVHYIMK 126
              |       |:|| :| |  |     : :   :|    :  :| :       |
Db        209 VAKEATEAAKCNLLAEKQYGFCKATLSEKLGGAEV---AVTCTVFQTQPVTSQ----PQP 261

Qy        127 NEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGT-PGKHVC-----GHHL 180
              ||  |              |   | |  | |||:    | | |: |  ||       || |
Db        262 EGANEAV------------PTPVVDPDAPPSPPLGAPGLPPAGSPPDSHVLLAAPPGHQL 309

Qy        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRE-SRPRRQGEVTVLSVG 226
              |           :  ||     :   : | | | ||:   |   |||
Db        310 HRAH--------YDLRHTFMGVVSLGSPSGEVSHPRKTRTVVQPSVG 348


RESULT 8
US-11-072-512-2203
; Sequence 2203, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2203
;   LENGTH: 241
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-2203

  Query Match             6.0%;  Score 85;  DB 7;  Length 241;
  Best Local Similarity   20.6%;  Pred. No. 1.1;
  Matches   40;  Conservative   38;  Mismatches   76;  Indels   40;  Gaps    9;

Qy         53 NGHPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEE-----EKV------ 101
              |   ||:|   :|  |    |    |:|           :: |  :|      :||
Db          4 NDAKEYLARREIPQLF-ESLLNGLMC----------SKPEDPVEYLESCLQKVKELGGC 51

Qy        102 EKIELNDSVNENSDTV-----GQIVHYIMKNEANADVLKA--MVADNSLYDPESPVTPST 154
              :|:: :  |::   |:     ||     ::||:: |: :   :: :  ::||  |
Db         52 DKVKWDTFVSQEKKTLPPLNGGQSRRSFLRNESDTDLSETAELIEEYEVFDPTRP----R 107

Qy        155 PGSPPVSPGPLSPGGTPGKHVC---GHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRE 211
              |    | || |  ||    :     |   :|| :: : :    :::|  |    : |
Db        108 PKIILVIGGPGSGKGTQSLKIAERYGFQYISVGELLRKKIHSTSSNRKWSLIAKIITTGE 167

Qy        212 SRPRRQGEVTVLSV 225
                |:   | |:  :
Db        168 LAPQ---ETTITEI 178


RESULT 9
US-11-240-769-67
; Sequence 67, Application US/11240769
; Publication No. US20060036089A1
; GENERAL INFORMATION:
;  APPLICANT: Soppet et al.
;  TITLE OF INVENTION: 33 Human Secreted Proteins
;  FILE REFERENCE: PZ037P1C2
;  CURRENT APPLICATION NUMBER: US/11/240,769
;  CURRENT FILING DATE:  2005-10-03
```

```
;  PRIOR APPLICATION NUMBER: 09/997,131
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 09/628,508
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: PCT/US00/03062
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: 60/119,468
;  PRIOR FILING DATE: 1999-02-10
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
;   LENGTH: 377
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (164)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (213)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-240-769-67

  Query Match             6.0%;  Score 85;  DB 7;  Length 377;
  Best Local Similarity   22.6%;  Pred. No. 2.1;
  Matches   53;  Conservative   28;  Mismatches   66;  Indels   88;  Gaps   14;

Qy          59 IAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVG 118
               |  ::| |||: |:  ::|   |:|           ||      |  :: |:: |:
Db         183 IINSVVVVFFLSGILSMIIIRTLRK-----------DIANYXKED-DIEDTMEESG---W 227

Qy         119 QIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGG--------T 170
               ::||         || :                  |||   |  | |
Db         228 KLVH--------GDVFR--------------------PPPVPHDPQLPAGLRHSAVLYD 258

Qy         171 PGKHVCGH------------HLHT-----VGGVVERDVCHRCRHKRW----HFIKPTNKS 209
               | :|:| |             | |:     | | | | |        |    || :|   |
Db         259 PHRHLCSHAWDAVALQPGSSHDHSLLPLHVHGGVWRIFC-------WPSVPHFKRP---S 308

Qy         210 RESRPRRQGEVTVLSVGRFRVTKVE--HKSNQKERRSLMSVSGAETVN--GEVPA 260
                | |    |  :|   | : : ::|  |       |   :|  |   |:   ::||
Db         309 VEERSLLYGN-SVPWCGFWHLLRIELLHLGKALIRSGALSHHGGSAVHVVRDLPA 362


RESULT 10
US-10-467-962B-31
; Sequence 31, Application US/10467962B
; Publication No. US20050246784A1
; GENERAL INFORMATION:
;  APPLICANT: Plesch, Gunnar
;  APPLICANT:  Blau, Astrid
;  APPLICANT:  Daeschner, Klaus
;  APPLICANT:  Klein, Mathieu
;  TITLE OF INVENTION: Identification of Herbicidally Active Substances
;  FILE REFERENCE: 2000_857
;  CURRENT APPLICATION NUMBER: US/10/467,962B
;  CURRENT FILING DATE:  2003-08-14
;  PRIOR APPLICATION NUMBER: PCT/EP02/01466
;  PRIOR FILING DATE: 2002-02-13
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 31
;   LENGTH: 954
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-467-962B-31

  Query Match             5.9%;  Score 84.5;  DB 6;  Length 954;
  Best Local Similarity   26.2%;  Pred. No. 8.4;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels   29;  Gaps    6;

Qy          77 ICHLLKKKGYRCTTEAEQDIEEEKVEKIELND----------SVNENSDTVG--QIVHY 123
               :|  || |    |   |  :|| :  : | :|           ::  : |||  | :
Db         785 LCSTLKVKAEDVTNRVENLLEELRAARKEASDLRSKAAVYKASVISNKAFTVGTSQTIRV 844

Qy         124 IMK--NEANADVLKAMVAD--NSLYDPESPVTPSTPGSPPV------SPGPLSPGGTPGK 173
               :::  :: :|| ||:      ::| || : |  |:|    |      ||| :| |   ||
Db         845 LVESMDDTDADSLKSAAEHLISTLEDPVAVVLGSSPEKDKVSLVAAFSPGVVSLGVQAGK 904

Qy         174 HV------CG 177
                :      ||
Db         905 FIGPIAKLCG 914


RESULT 11
US-11-072-512-2452
; Sequence 2452, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2452
;   LENGTH: 1344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-2452

  Query Match             5.9%;  Score 84;  DB 7;  Length 1344;
  Best Local Similarity   24.9%;  Pred. No. 15;
  Matches   53;  Conservative   26;  Mismatches   74;  Indels   60;  Gaps   12;

Qy          11 VLAAAV-FVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGH-PEYIAYALVPVFF 68
               |||  :  : | : |||:                |||       |:   ||     |:|
Db         633 VLATTLNMLPGLIHSPLIC-------------TTPKHYIRFGSPPIPERRRRPLLP--- 675

Qy          69 IMGLFGVLICHLLK---KKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQIVHYIM 125
                 | |        :|   ::|   |:   |:   : : :          ||||  :       |
Db         676 -DGTFSSCKKRWIKQALEEGMTQTSSVPQETRTQHLYQ------SNENSSSSS-----IC 723

Qy         126 KNEANADVLKAMVADNSLY----------DPESPVTPSTPGSPPVSPGPLSPGGT-PGKH 174
               |:  |||:|  :    | |           | ||||| | |   ||   :|| : ||:
Db         724 KD--NADLLSPLKKWKSRYLMEQNVTKLLRPLSPVTPPPPNSGSKSPQLATPGSSHPGEE 781

Qy         175 VC--GHHL-----------HTVGGVVERDVCHR 194
                 |  |: |                 :: ::|||
Db         782 ECRNGYSLMFSPVTSLTTASRCNTPLQFELCHR 814


RESULT 12
```

```
US-10-467-657-1672
; Sequence 1672, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
;  APPLICANT: CHIRON SpA
;  APPLICANT: FONTANA Maria Rita
;  APPLICANT: PIZZA Mariagrazia
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MONACI Elisabetta
;  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/467,657
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: GB-0103424.8
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1672
;   LENGTH: 399
;   TYPE: PRT
;   ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1672

  Query Match             5.8%;  Score 82.5;  DB 6;  Length 399;
  Best Local Similarity   25.4%;  Pred. No. 3.9;
  Matches   34;  Conservative   15;  Mismatches   58;  Indels   27;  Gaps    4;

Qy         52 GNGHPEYIAYALVPVFFIM----GLFGVLICH----------LLKKKGYRCTTEAEQDI 96
              | | ||   ||:  |  :|     :||: : |          |  :  :       ||:
Db        260 GPGDPEPCTYAIEAVQKLMESGKPIFGICLGHQLISLAIGAKTLKMRFSHHGANHPVQDL 319

Qy         97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTP--ST 154
              :  ||      |     ::||:           ||| :    : ||:|   |     |
Db        320 DSGKVVITSQNHGFAVDADTL----------PANARITHKSLFDNTLQGIELTDKPVFCF 369

Qy        155 PGSPPVSPGPLSPG 168
              | |  ||||   |
Db        370 QGHPEASPGPQDVG 383


RESULT 13
US-11-169-232-104
; Sequence 104, Application US/11169232
; Publication No. US20060025570A1
;  GENERAL INFORMATION:
;       APPLICANT: Turner, Katherine
;                  Clark, Stephen C.
;                  Jacobs, Kenneth
;                  Hewick, Rodney M.
;                  Gesner, Thomas G.
;       TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;       NUMBER OF SEQUENCES: 143
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/11/169,232
;            FILING DATE: 28-Jun-2005
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/10/124,557
;            FILING DATE: 16-Apr-2002
;            APPLICATION NUMBER: US 07/643,502
;            FILING DATE: 18-JAN-1991
;            APPLICATION NUMBER: US 07/546,114
;            FILING DATE: 29-JUN-1990
;            APPLICATION NUMBER: US 07/457,196
;            FILING DATE: 29-DEC-1989
;            APPLICATION NUMBER: US 07/390,901
;            FILING DATE: 08-AUG-1989
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Cserr, Luann
;            REGISTRATION NUMBER: 31,822
;            REFERENCE/DOCKET NUMBER: GI 5190
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)876-1170
;            TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 104:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1140 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-11-169-232-104

  Query Match             5.8%;  Score 82.5;  DB 7;  Length 1140;
  Best Local Similarity   20.1%;  Pred. No. 17;
  Matches   55;  Conservative   35;  Mismatches  111;  Indels   73;  Gaps    9;

Qy         23 SSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIAYALVPVFFIMGLFGVLICHLLK 82
              ||     | :|:|:|: | |: :| | |
Db        115 SSKKAPPPSGASQTIKSTTKRSPKPPN-------------------------------- 141

Qy         83 KKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI--VHYIMKNEANADVLKAM-VA 139
              ||  :   |:|:  ||  | : : : | : :| :   |  :      : || :: | : |
Db        142 KKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVK 201

Qy        140 DNSLYDPESPVTPSTPGSPPVSPGPLSPGG----TPGKHVCGHHLHTVGGVVERDVCHRC 195
              ||      :   ||  |        | |  |      ||       |:  :    :
Db        202 DNKKNRTKKKPTPKPPVVDEAGSG-LDNGDFKVTTPDTSTTQHNKVSTSPKITT------ 254

Qy        196 RHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHK----------------SN 238
                  || |  | | |          |:   : | || |                  ::
Db        255 -------AKPIN-PRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTS 306

Qy        239 QKERRSLMSVSGAETV-NGEVPATPVKRERSGTE 271
              || :|:   |  :     :| | |  :  : |:
Db        307 AKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTK 340


RESULT 14
US-11-169-232-52
; Sequence 52, Application US/11169232
; Publication No. US20060025570A1
;  GENERAL INFORMATION:
;       APPLICANT: Turner, Katherine
;                  Clark, Stephen C.
;                  Jacobs, Kenneth
;                  Hewick, Rodney M.
;                  Gesner, Thomas G.
;       TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;       NUMBER OF SEQUENCES: 143
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/11/169,232
```

```
;             FILING DATE: 28-Jun-2005
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 52:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1363 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-11-169-232-52

  Query Match             5.8%;  Score 82.5;  DB 7;  Length 1363;
  Best Local Similarity   20.1%;  Pred. No. 21;
  Matches   55;  Conservative   35;  Mismatches  111;  Indels   73;  Gaps    9;

Qy         23 SSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIAYALVPVFFIMGLFGVLICHLLK 82
              ||     | :|:|:|: | |: :| | |
Db         74 SSKKAPPPSGASQTIKSTTKRSPKPPN---------------------------------- 100

Qy         83 KKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI--VHYIMKNEANADVLKAM-VA 139
              ||  :   |:|:  ||  | : : : | : :| :   |  :     : || :: | : |
Db        101 KKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVK 160

Qy        140 DNSLYDPESPVTPSTPGSPPVSPGPLSPGG----TPGKHVCGHHLHTVGGVVERDVCHRC 195
              ||      :   ||  |       | |  |      ||      |:  :    :
Db        161 DNKKNRTKKKPTPKPPVVDEAGSG-LDNGDFKVTTPDTSTTQHNKVSTSPKITT------ 213

Qy        196 RHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHK-----------------SN 238
                  || |  | | |          |:   : | || |                    ::
Db        214 -------AKPIN-PRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTS 265

Qy        239 QKERRSLMSVSGAETV-NGEVPATPVKRERSGTE 271
               || :|:   |  :     :| | |  :  : |:
Db        266 AKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTK 299


RESULT 15
US-11-169-232-2
; Sequence 2, Application US/11169232
; Publication No. US20060025570A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/11/169,232
;             FILING DATE: 28-Jun-2005
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990\
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1404 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-11-169-232-2

  Query Match             5.8%;  Score 82.5;  DB 7;  Length 1404;
  Best Local Similarity   20.1%;  Pred. No. 22;
  Matches   55;  Conservative   35;  Mismatches  111;  Indels   73;  Gaps    9;

Qy         23 SSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIAYALVPVFFIMGLFGVLICHLLK 82
              ||     | :|:|:|: | |: :| | |
Db        115 SSKKAPPPSGASQTIKSTTKRSPKPPN---------------------------------- 141

Qy         83 KKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI--VHYIMKNEANADVLKAM-VA 139
              ||  :   |:|:  ||  | : : : | : :| :   |  :     : || :: | : |
Db        142 KKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVK 201

Qy        140 DNSLYDPESPVTPSTPGSPPVSPGPLSPGG----TPGKHVCGHHLHTVGGVVERDVCHRC 195
              ||      :   ||  |       | |  |      ||      |:  :    :
Db        202 DNKKNRTKKKPTPKPPVVDEAGSG-LDNGDFKVTTPDTSTTQHNKVSTSPKITT------ 254

Qy        196 RHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHK-----------------SN 238
                  || |  | | |      .    |:   : | || |                    ::
Db        255 -------AKPIN-PRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTS 306

Qy        239 QKERRSLMSVSGAETV-NGEVPATPVKRERSGTE 271
               || :|:   |  :     :| | |  :  : |:
Db        307 AKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTK 340


Search completed: February 23, 2006, 12:12:39
Job time : 19 secs
```

OM protein - protein search, using sw model

Run on: February 23, 2006, 12:08:48 ; Search time 163 Seconds
(without alignments)
694.673 Million cell updates/sec

Title: US-10-620-562-2
Perfect score: 1426
Sequence: 1 MAPRALPGSAVLAAAVFVGG.....ETVNGEVPATPVKRERSGTE 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1426	100.0	271	3	US-09-910-562-2	Sequence 2, Appli
2	1426	100.0	271	5	US-10-620-562-2	Sequence 2, Appli
3	477.5	33.5	103	3	US-09-933-767-294	Sequence 294, App
4	477.5	33.5	103	4	US-10-004-860-294	Sequence 294, App
5	477.5	33.5	103	4	US-10-023-282-294	Sequence 294, App
6	332	23.3	181	3	US-09-910-562-6	Sequence 6, Appli
7	332	23.3	181	3	US-09-933-767-633	Sequence 633, App
8	332	23.3	181	4	US-10-004-860-633	Sequence 633, App
9	332	23.3	181	4	US-10-023-282-633	Sequence 633, App
10	332	23.3	181	5	US-10-620-562-6	Sequence 6, Appli
11	307	21.5	408	3	US-09-057-951-4	Sequence 4, Appli
12	307	21.5	408	4	US-10-105-150-4	Sequence 4, Appli
13	307	21.5	430	3	US-09-057-951-2	Sequence 2, Appli
14	307	21.5	430	3	US-09-836-607-2	Sequence 2, Appli
15	307	21.5	430	3	US-09-421-112-2	Sequence 2, Appli
16	307	21.5	430	4	US-10-105-150-2	Sequence 2, Appli
17	307	21.5	430	4	US-10-146-574-2	Sequence 2, Appli
18	307	21.5	430	4	US-10-157-031-265	Sequence 265, App
19	307	21.5	430	4	US-10-322-281-522	Sequence 522, App
20	295.5	20.7	436	4	US-10-146-574-4	Sequence 4, Appli
21	289.5	20.3	357	4	US-10-322-281-515	Sequence 515, App
22	281	19.7	294	3	US-09-764-875-680	Sequence 680, App
23	278.5	19.5	194	3	US-09-910-562-4	Sequence 4, Appli
24	278.5	19.5	194	5	US-10-620-562-4	Sequence 4, Appli
25	194	13.6	250	4	US-10-322-281-520	Sequence 520, App
26	193.5	13.6	299	4	US-10-094-749-2913	Sequence 2913, Ap
27	147	10.3	29	3	US-09-910-562-7	Sequence 7, Appli
28	147	10.3	29	3	US-09-933-767-634	Sequence 634, App
29	147	10.3	29	4	US-10-004-860-634	Sequence 634, App
30	147	10.3	29	4	US-10-023-282-634	Sequence 634, App
31	147	10.3	29	5	US-10-620-562-7	Sequence 7, Appli
32	131	9.2	66	3	US-09-796-692-1534	Sequence 1534, Ap
33	131	9.2	66	4	US-10-040-862-1534	Sequence 1534, Ap
34	131	9.2	66	4	US-10-057-475B-1534	Sequence 1534, Ap
35	131	9.2	66	4	US-10-154-884B-1534	Sequence 1534, Ap
36	131	9.2	66	4	US-10-764-324-1534	Sequence 1534, Ap
37	114	8.0	96	4	US-10-322-281-518	Sequence 518, App
38	102.5	7.2	1829	4	US-10-437-963-192144	Sequence 192144,
39	97.5	6.8	447	4	US-10-104-047-2688	Sequence 2688, Ap
40	97.5	6.8	1049	5	US-10-487-561-80	Sequence 80, Appl
41	97.5	6.8	1513	4	US-10-369-493-19093	Sequence 19093, A
42	97.5	6.8	1746	4	US-10-220-955-17	Sequence 17, Appl
43	97	6.8	287	4	US-10-290-631-6	Sequence 6, Appli
44	97	6.8	287	4	US-10-290-631-8	Sequence 8, Appli
45	97	6.8	287	4	US-10-777-524-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-910-562-2
; Sequence 2, Application US/09910562
; Patent No. US20020098163A1
; GENERAL INFORMATION:
; APPLICANT: Zeng et al.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptors TR21 and TR22
; FILE REFERENCE: PF530
; CURRENT APPLICATION NUMBER: US/09/910,562
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 60/221,143
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,116
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 271
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-910-562-2

Query Match 100.0%; Score 1426; DB 3; Length 271;
Best Local Similarity 100.0%; Pred. No. 4.4e-118;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180

Qy        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240

Qy        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
              ||||||||||||||||||||||||||||||
Db        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
```

RESULT 2
US-10-620-562-2
; Sequence 2, Application US/10620562

```
; Publication No. US20050090436A1
; GENERAL INFORMATION:
;  APPLICANT: Zeng, Z et al.
;  TITLE OF INVENTION: Human Tumor Necrosis Factor receptors TR21 and TR22
;  FILE REFERENCE: PF530C1
;  CURRENT APPLICATION NUMBER: US/10/620,562
;  CURRENT FILING DATE:  2003-07-17
;  PRIOR APPLICATION NUMBER: 09/910,562
;  PRIOR FILING DATE: 2001-07-23
;  PRIOR APPLICATION NUMBER: 60/221,143
;  PRIOR FILING DATE: 2000-07-27
;  PRIOR APPLICATION NUMBER: 60/220,116
;  PRIOR FILING DATE: 2000-07-24
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-620-562-2

  Query Match            100.0%;  Score 1426;  DB 5;  Length 271;
  Best Local Similarity  100.0%;  Pred. No. 4.4e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLICHLLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLICHLLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180

Qy        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240

Qy        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
              |||||||||||||||||||||||||||||||
Db        241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271


RESULT 3
US-09-933-767-294
; Sequence 294, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P2
;  CURRENT APPLICATION NUMBER: US/09/933,767
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: PCT/US01/05614
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: 60/184,836
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/193,170
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 09/205,258
;  PRIOR FILING DATE: 1998-12-04
;  PRIOR APPLICATION NUMBER: PCT/US98/11422
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/048,885
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,375
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,881
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,880
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,896
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,020
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,876
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,895
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,884
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,894
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,971
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,964
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,882
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,899
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,893
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,900
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,901
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,892
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,915
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,019
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,970
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,972
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,916
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,373
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,875
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,374
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,917
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,949
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,974
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,883
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,897
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,898
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,962
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,963
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,877
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,878
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/068,054
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,064
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,053
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/070,923
;  PRIOR FILING DATE: 1997-12-18
```

```
;  PRIOR APPLICATION NUMBER: 60/073,160
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/073,159
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/073,165
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/073,164
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/085,925
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,921
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,923
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,922
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/092,921
;  PRIOR FILING DATE: 1998-07-15
;  PRIOR APPLICATION NUMBER: 60/094,657
;  PRIOR FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1245
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 294
;   LENGTH: 103
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (78)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (80)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (81)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (82)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (103)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-933-767-294

  Query Match            33.5%;  Score 477.5;  DB 3;  Length 103;
  Best Local Similarity  95.0%;  Pred. No. 2e-34;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVPFFIMGLFGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||     ||||||||||||||||||||
Db         61 YALVPVPFFIMGLFGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101


RESULT 4
US-10-004-860-294
; Sequence 294, Application US/10004860
; Publication No. US20030065160A1
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/004,860
;  CURRENT FILING DATE:  2001-12-07
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 294
;   LENGTH: 103
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (78)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (80)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (81)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (82)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (103)
;   OTHER INFORMATION: Xaa equals stop translation
US-10-004-860-294

  Query Match            33.5%;  Score 477.5;  DB 4;  Length 103;
  Best Local Similarity  95.0%;  Pred. No. 2e-34;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVPFFIMGLFGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||     ||||||||||||||||||||
Db         61 YALVPVPFFIMGLFGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101


RESULT 5
US-10-023-282-294
; Sequence 294, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/023,282
;  CURRENT FILING DATE: 2001-12-20
;  EARLIER APPLICATION NUMBER: 09/205,258
;  EARLIER FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
```

```
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 294
;   LENGTH: 103
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (78)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (80)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (81)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (82)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (103)
;   OTHER INFORMATION: Xaa equals stop translation
US-10-023-282-294

  Query Match             33.5%;  Score 477.5;  DB 4;  Length 103;
  Best Local Similarity   95.0%;  Pred. No. 2e-34;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||  ||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101


RESULT 6
US-09-910-562-6
; Sequence 6, Application US/09910562
; Patent No. US20020098163A1
; GENERAL INFORMATION:
;  APPLICANT: Zeng et al.
;  TITLE OF INVENTION: Human Tumor Necrosis Factor Receptors TR21 and TR22
;  FILE REFERENCE: PF530
;  CURRENT APPLICATION NUMBER: US/09/910,562
;  CURRENT FILING DATE:  2001-07-23
;  PRIOR APPLICATION NUMBER: 60/221,143
;  PRIOR FILING DATE: 2000-07-27
;  PRIOR APPLICATION NUMBER: 60/220,116
;  PRIOR FILING DATE: 2000-07-24
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 181
;   TYPE: PRT
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (56)..(57)
;   OTHER INFORMATION: Xaa equals any amino acid
;   NAME/KEY: SITE
;   LOCATION: (60)..(60)
;   OTHER INFORMATION: Xaa equals any amino acid
;   NAME/KEY: SITE
;   LOCATION: (83)..(84)
;   OTHER INFORMATION: Xaa equals any amino acid
;   NAME/KEY: SITE
;   LOCATION: (165)..(165)
;   OTHER INFORMATION: Xaa equals any amino acid
US-09-910-562-6

  Query Match             23.3%;  Score 332;  DB 3;  Length 181;
  Best Local Similarity   85.5%;  Pred. No. 3.6e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
              :::||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy        157 SPPVSPGPLSPGGTPG 172
              |||||||    ||  |
Db        136 SPPVSPGLCHQGGRQG 151


RESULT 7
US-09-933-767-633
```

; Sequence 633, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: PCT/US01/05614
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,836
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/193,170
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/205,258
; PRIOR FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: PCT/US98/11422
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/048,885
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,375
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,881
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,880
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,896
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,020
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,876
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,895
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,884
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,894
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,971
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,882
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,899
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,893
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,900
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,901
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,892
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,915
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,019
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,972
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,916
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,373
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,875
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,374
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,917
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,949
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,883
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,897
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,898
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,962
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,963
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,877
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,878
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 633
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (56)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (57)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (60)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (165)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-933-767-633

Query Match 23.3%; Score 332; DB 3; Length 181;

```
  Best Local Similarity   85.5%;  Pred. No. 3.6e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
               :::|||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy         157 SPPVSPGPLSPGGTPG 172
               ||||||| || |
Db         136 SPPVSPGLCHQGGRQG 151


RESULT 8
US-10-004-860-633
; Sequence 633, Application US/10004860
; Publication No. US20030065160A1
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/004,860
;  CURRENT FILING DATE:  2001-12-07
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 633
;   LENGTH: 181
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (56)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (57)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (60)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (83)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (84)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (165)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-004-860-633

  Query Match             23.3%;  Score 332;  DB 4;  Length 181;
  Best Local Similarity   85.5%;  Pred. No. 3.6e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
               :::|||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy         157 SPPVSPGPLSPGGTPG 172
               ||||||| || |
Db         136 SPPVSPGLCHQGGRQG 151


RESULT 9
US-10-023-282-633
; Sequence 633, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/023,282
;  CURRENT FILING DATE: 2001-12-20
;  EARLIER APPLICATION NUMBER: 09/205,258
;  EARLIER FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
```

```
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 633
;   LENGTH: 181
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (56)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (57)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (60)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (83)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (84)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (165)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-023-282-633

  Query Match             23.3%;  Score 332;  DB 4;  Length 181;
  Best Local Similarity   85.5%;  Pred. No. 3.6e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
               :::|||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy         157 SPPVSPGPLSPGGTPG 172
               |||||||    ||  |
Db         136 SPPVSPGLCHQGGRQG 151


RESULT 10
US-10-620-562-6
; Sequence 6, Application US/10620562
; Publication No. US20050090436A1
; GENERAL INFORMATION:
;  APPLICANT: Zeng, Z et al.
;  TITLE OF INVENTION: Human Tumor Necrosis Factor receptors TR21 and TR22
;  FILE REFERENCE: PF530C1
;  CURRENT APPLICATION NUMBER: US/10/620,562
;  CURRENT FILING DATE:  2003-07-17
;  PRIOR APPLICATION NUMBER: 09/910,562
;  PRIOR FILING DATE: 2001-07-23
;  PRIOR APPLICATION NUMBER: 60/221,143
;  PRIOR FILING DATE: 2000-07-27
;  PRIOR APPLICATION NUMBER: 60/220,116
;  PRIOR FILING DATE: 2000-07-24
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 181
;   TYPE: PRT
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (56)..(57)
;   OTHER INFORMATION: X equals any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (60)..(60)
;   OTHER INFORMATION: X equals any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (83)..(84)
;   OTHER INFORMATION: X equals any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (165)..(165)
;   OTHER INFORMATION: X equals any amino acid
US-10-620-562-6

  Query Match             23.3%;  Score 332;  DB 5;  Length 181;
  Best Local Similarity   85.5%;  Pred. No. 3.6e-21;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
               :::|||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy         157 SPPVSPGPLSPGGTPG 172
               |||||||    ||  |
Db         136 SPPVSPGLCHQGGRQG 151


RESULT 11
US-09-057-951-4
; Sequence 4, Application US/09057951
; Patent No. US20020025551A1
; GENERAL INFORMATION:
;   APPLICANT:  Holtzman, Douglas
;   TITLE OF INVENTION:  NOVEL MOLECULES OF THE T129-RELATED
;   TITLE OF INVENTION:  PROTEIN FAMILY AND USES THEREOF
;   NUMBER OF SEQUENCES:  8
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Fish & Richardson P.C.
;     STREET:  225 Franklin Street
;     CITY:  Boston
;     STATE:  MA
;     COUNTRY:  USA
;     ZIP:  02110-2804
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Diskette
;     COMPUTER:  IBM Compatible
;     OPERATING SYSTEM:  Windows95
;     SOFTWARE:  FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/057,951
;     FILING DATE:  09-APR-1998
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Meiklejohn, Ph.D., Anita L.
;     REGISTRATION NUMBER:  35,283
;     REFERENCE/DOCKET NUMBER:  09404/046001
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  617/542-5070
;     TELEFAX:  617/542-8906
;     TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  4:
;   SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  408 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-057-951-4

  Query Match             21.5%;  Score 307;  DB 3;  Length 408;
  Best Local Similarity   31.0%;  Pred. No. 1.8e-18;
  Matches   84;  Conservative   36;  Mismatches   69;  Indels   82;  Gaps   12;

Qy         44 TPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
              |  | | |  | ||      |   |:|||| :||| |:|:|:|||:||| ||   |
Db        122 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 181

Qy         94 ------------QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
                        :|                 | ||:| :|  | : : ||  |: :: :
Db        182 GGGGSGINPAYRTED---------------ANEDTIGVLVRLITEKKENAAALEELLKE 225

Qy        141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCR 196
              |  :  | |     |||  | :|   |  |:|   |||||| |:  :    | ||
Db        226 ---YHSKQLVQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCS 277

Qy        197 HKRWHFI-----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKS 237
              |:|  :                 ||   :  ::  ||||:|:||||||| ::
Db        278 QKKWPEVLLSPEAVAATTPVPSLLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI---- 333

Qy        238 NQKERRSLMSVSGAETVN------GEVPATP 262
                |:|:   ||  :|:       |::| :|
Db        334 --PEQRTSSMVSEVKTITEAGPSWGDLPDSP 362


RESULT 12
US-10-105-150-4
; Sequence 4, Application US/10105150
; Publication No. US20020119524A1
;   GENERAL INFORMATION:
;        APPLICANT: Holtzman, Douglas
;        TITLE OF INVENTION: NOVEL MOLECULES OF THE T129-RELATED
;                            PROTEIN FAMILY AND USES THEREOF
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows95
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/105,150
;             FILING DATE: 25-Mar-2002
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/057,951
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Meiklejohn, Ph.D., Anita L.
;             REGISTRATION NUMBER: 35,283
;             REFERENCE/DOCKET NUMBER: 09404/046001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 408 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-105-150-4

  Query Match             21.5%;  Score 307;  DB 4;  Length 408;
  Best Local Similarity   31.0%;  Pred. No. 1.8e-18;
  Matches   84;  Conservative   36;  Mismatches   69;  Indels   82;  Gaps   12;

Qy         44 TPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
              |  | | |  | ||      |   |:|||| :||| |:|:|:|||:||| ||   |
Db        122 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 181

Qy         94 ------------QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
                        :|                 | ||:| :|  | : : ||  |: :: :
Db        182 GGGGSGINPAYRTED---------------ANEDTIGVLVRLITEKKENAAALEELLKE 225

Qy        141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCR 196
              |  :  | |     |||  | :|   |  |:|   |||||| |:  :    | ||
Db        226 ---YHSKQLVQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCS 277

Qy        197 HKRWHFI-----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKS 237
              |:|  :                 ||   :  ::  ||||:|:||||||| ::
Db        278 QKKWPEVLLSPEAVAATTPVPSLLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI---- 333

Qy        238 NQKERRSLMSVSGAETVN------GEVPATP 262
                |:|:   ||  :|:       |::| :|
Db        334 --PEQRTSSMVSEVKTITEAGPSWGDLPDSP 362


RESULT 13
US-09-057-951-2
; Sequence 2, Application US/09057951
; Patent No. US20020025551A1
;  GENERAL INFORMATION:
;    APPLICANT:  Holtzman, Douglas
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE T129-RELATED
;    TITLE OF INVENTION:  PROTEIN FAMILY AND USES THEREOF
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/057,951
;      FILING DATE:  09-APR-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  09404/046001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  430 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-057-951-2

  Query Match             21.5%;  Score 307;  DB 3;  Length 430;
  Best Local Similarity   31.0%;  Pred. No. 1.9e-18;
  Matches   84;  Conservative   36;  Mismatches   69;  Indels   82;  Gaps   12;

Qy         44 TPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
```

```
                |  | | |  | ||     |   |:|||| :||| |:|:|:|||:||| ||   |
Db        144 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 203

Qy         94 ------------QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
                          :|                  | ||:| :|  | : : ||  |: :: :
Db        204 GGGGSGINPAYRTED---------------ANEDTIGVLVRLITEKKENAAALEELLKE 247

Qy        141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCR 196
                 |  :  | |     |||  | :|   |  |:|   |||||| |:  :    | ||
Db        248 ---YHSKQLVQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCS 299

Qy        197 HKRWHFI-----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKS 237
                |:|  :                  ||   :  ::  ||||:|:||||||| ::
Db        300 QKKWPEVLLSPEAVAATTPVPSLLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI---- 355

Qy        238 NQKERRSLMSVSGAETVN------GEVPATP 262
                |:|:   ||  :|:       |::| :|
Db        356 --PEQRTSSMVSEVKTITEAGPSWGDLPDSP 384


RESULT 14
US-09-836-607-2
; Sequence 2, Application US/09836607
; Patent No. US20020098541A1
; GENERAL INFORMATION:
;  APPLICANT: Ni, Jian
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: TNFR Related Gene 12
;  FILE REFERENCE: PF490P1
;  CURRENT APPLICATION NUMBER: US/09/836,607
;  CURRENT FILING DATE:  2001-04-17
;  PRIOR APPLICATION NUMBER: 60/198,388
;  PRIOR FILING DATE: 2000-04-19
;  PRIOR APPLICATION NUMBER: 09/421,112
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/104,950
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 430
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-836-607-2

  Query Match             21.5%;  Score 307;  DB 3;  Length 430;
  Best Local Similarity   31.0%;  Pred. No. 1.9e-18;
  Matches   84;  Conservative   36;  Mismatches   69;  Indels   82;  Gaps   12;

Qy         44 TPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
                |  | | |  | ||     |   |:|||| :||| |:|:|:|||:||| ||   |
Db        144 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 203

Qy         94 ------------QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
                          :|                  | ||:| :|  | : : ||  |: :: :
Db        204 GGGGSGINPAYRTED---------------ANEDTIGVLVRLITEKKENAAALEELLKE 247

Qy        141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCR 196
                 |  :  | |     |||  | :|   |  |:|   |||||| |:  :    | ||
Db        248 ---YHSKQLVQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCS 299

Qy        197 HKRWHFI-----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKS 237
                |:|  :                  ||   :  ::  ||||:|:||||||| ::
Db        300 QKKWPEVLLSPEAVAATTPVPSLLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI---- 355

Qy        238 NQKERRSLMSVSGAETVN------GEVPATP 262
                |:|:   ||  :|:       |::| :|
Db        356 --PEQRTSSMVSEVKTITEAGPSWGDLPDSP 384


RESULT 15
US-09-421-112-2
; Sequence 2, Application US/09421112
; Publication No. US20030082532A1
; GENERAL INFORMATION:
;  APPLICANT: Ni, Jian
;  APPLICANT: Ruben, Steven M.
;  TITLE OF INVENTION: TNFR Related Gene 12
;  FILE REFERENCE: PF490
;  CURRENT APPLICATION NUMBER: US/09/421,112
;  CURRENT FILING DATE: 1999-10-19
;  EARLIER APPLICATION NUMBER: 60/104,950
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 430
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-421-112-2

  Query Match             21.5%;  Score 307;  DB 3;  Length 430;
  Best Local Similarity   31.0%;  Pred. No. 1.9e-18;
  Matches   84;  Conservative   36;  Mismatches   69;  Indels   82;  Gaps   12;

Qy         44 TPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
                |  | | |  | ||     |   |:|||| :||| |:|:|:|||:||| ||   |
Db        144 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 203

Qy         94 ------------QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
                          :|                  | ||:| :|  | : : ||  |: :: :
Db        204 GGGGSGINPAYRTED---------------ANEDTIGVLVRLITEKKENAAALEELLKE 247

Qy        141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCR 196
                 |  :  | |     |||  | :|   |  |:|   |||||| |:  :    | ||
Db        248 ---YHSKQLVQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCS 299

Qy        197 HKRWHFI-----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKS 237
                |:|  :                  ||   :  ::  ||||:|:||||||| ::
Db        300 QKKWPEVLLSPEAVAATTPVPSLLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI---- 355

Qy        238 NQKERRSLMSVSGAETVN------GEVPATP 262
                |:|:   ||  :|:       |::| :|
Db        356 --PEQRTSSMVSEVKTITEAGPSWGDLPDSP 384


Search completed: February 23, 2006, 12:12:15
Job time : 164 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 23, 2006, 12:08:03 ; Search time 47 Seconds
(without alignments)
476.704 Million cell updates/sec

Title: US-10-620-562-2
Perfect score: 1426
Sequence: 1 MAPRALPGSAVLAAAVFVGG.....ETVNGEVPATPVKRERSGTE 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	477.5	33.5	103	2	US-09-205-258-294	Sequence 294, App
2	477.5	33.5	103	2	US-10-004-860-294	Sequence 294, App
3	332	23.3	181	2	US-09-205-258-633	Sequence 633, App
4	332	23.3	181	2	US-10-004-860-633	Sequence 633, App
5	282	19.8	518	2	US-09-949-016-10986	Sequence 10986, A
6	147	10.3	29	2	US-09-205-258-634	Sequence 634, App
7	147	10.3	29	2	US-10-004-860-634	Sequence 634, App
8	103.5	7.3	819	2	US-09-248-796A-15758	Sequence 15758, A
9	97.5	6.8	447	2	US-10-104-047-2688	Sequence 2688, Ap
10	97	6.8	287	2	US-08-985-950-6	Sequence 6, Appli
11	97	6.8	287	2	US-08-985-950-8	Sequence 8, Appli
12	97	6.8	287	2	US-09-546-049-6	Sequence 6, Appli
13	97	6.8	287	2	US-09-546-049-8	Sequence 8, Appli
14	95	6.7	992	2	US-08-890-865A-1	Sequence 1, Appli
15	94	6.6	793	1	US-08-015-985-3	Sequence 3, Appli
16	94	6.6	793	2	US-09-280-597-3	Sequence 3, Appli
17	92.5	6.5	591	2	US-09-248-796A-14242	Sequence 14242, A
18	91.5	6.4	527	2	US-09-370-838-216	Sequence 216, App
19	91.5	6.4	527	2	US-09-854-133-216	Sequence 216, App
20	90.5	6.3	8991	2	US-08-714-741-32	Sequence 32, Appl
21	89.5	6.3	901	2	US-09-828-062-8	Sequence 8, Appli
22	89	6.2	401	1	US-08-549-004A-5	Sequence 5, Appli
23	89	6.2	401	2	US-09-051-982A-5	Sequence 5, Appli
24	89	6.2	605	2	US-09-487-558B-428	Sequence 428, App
25	88	6.2	330	2	US-09-270-767-41611	Sequence 41611, A
26	88	6.2	539	2	US-08-906-360-1	Sequence 1, Appli
27	87.5	6.1	559	2	US-08-956-171E-5251	Sequence 5251, Ap
28	87.5	6.1	559	2	US-08-781-986A-5251	Sequence 5251, Ap
29	87.5	6.1	1420	1	US-08-540-804-14	Sequence 14, Appl
30	87.5	6.1	1420	1	US-08-218-265-14	Sequence 14, Appl
31	87.5	6.1	1420	2	US-08-521-872-14	Sequence 14, Appl
32	87.5	6.1	1420	2	US-08-590-399-14	Sequence 14, Appl
33	87.5	6.1	1420	2	US-09-487-558B-356	Sequence 356, App
34	87	6.1	761	2	US-09-198-452A-335	Sequence 335, App
35	87	6.1	902	2	US-09-438-185A-319	Sequence 319, App
36	86.5	6.1	290	2	US-09-270-767-35246	Sequence 35246, A
37	86.5	6.1	290	2	US-09-270-767-50463	Sequence 50463, A
38	86	6.0	204	2	US-08-529-055-51	Sequence 51, Appl
39	86	6.0	547	2	US-09-221-013A-14	Sequence 14, Appl
40	86	6.0	729	2	US-09-949-016-11190	Sequence 11190, A
41	86	6.0	2753	2	US-09-949-016-7659	Sequence 7659, Ap
42	86	6.0	2753	2	US-09-949-016-7660	Sequence 7660, Ap
43	86	6.0	3924	2	US-09-538-092-1246	Sequence 1246, Ap
44	85.5	6.0	168	2	US-08-529-055-55	Sequence 55, Appl
45	85.5	6.0	183	2	US-08-529-055-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
US-09-205-258-294
; Sequence 294, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06

```
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 294
;   LENGTH: 103
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (78)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (80)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (81)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (82)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (103)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-294

  Query Match             33.5%;  Score 477.5;  DB 2;  Length 103;
  Best Local Similarity   95.0%;  Pred. No. 1.3e-40;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||  ||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101


RESULT 2
US-10-004-860-294
; Sequence 294, Application US/10004860
; Patent No. 6914047
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/004,860
;  CURRENT FILING DATE:  2001-12-07
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 294
;   LENGTH: 103
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (78)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (80)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (81)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (82)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (103)
;   OTHER INFORMATION: Xaa equals stop translation
US-10-004-860-294

  Query Match             33.5%;  Score 477.5;  DB 2;  Length 103;
  Best Local Similarity   95.0%;  Pred. No. 1.3e-40;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

Qy         61 YALVPVFFIMGLFGVLIC-HLLKKKGYRCTTEAEQDIEEEK 100
              |||||||||||||||||  ||||||||||||||||||
Db         61 YALVPVFFIMGLFGVLIXPXXXKKKGYRCTTEAEQDIEEEK 101


RESULT 3
US-09-205-258-633
; Sequence 633, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/09/205,258
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 633
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (56)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (57)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (60)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (165)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-633

Query Match 23.3%; Score 332; DB 2; Length 181;
Best Local Similarity 85.5%; Pred. No. 1.6e-25;
Matches 65; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

```
Qy          97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
               :::||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy         157 SPPVSPGPLSPGGTPG 172
               |||||||    ||  |
Db         136 SPPVSPGLCHQGGRQG 151
```

RESULT 4
US-10-004-860-633
; Sequence 633, Application US/10004860
; Patent No. 6914047
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/10/004,860
; CURRENT FILING DATE: 2001-12-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 633
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (56)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (57)

```
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (60)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (83)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (84)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (165)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-004-860-633

  Query Match            23.3%;  Score 332;  DB 2;  Length 181;
  Best Local Similarity  85.5%;  Pred. No. 1.6e-25;
  Matches   65;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 156
              :::||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 KKKKVEKXXLNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPG 135

Qy        157 SPPVSPGPLSPGGTPG 172
              |||||||    ||  |
Db        136 SPPVSPGLCHQGGRQG 151


RESULT 5
US-09-949-016-10986
; Sequence 10986, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10986
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10986

  Query Match            19.8%;  Score 282;  DB 2;  Length 518;
  Best Local Similarity  36.9%;  Pred. No. 8.4e-20;
  Matches   80;  Conservative   29;  Mismatches   84;  Indels   24;  Gaps    8;

Qy         55 HPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENS 114
              |  |: : || |||:||| | :|||:|||||||| |    : :: :::  | :| :||
Db        265 HGLYMLFLLVLVFFLMGLVGFMICHVLKKKGYRCRTSRGSEPDDAQLQPPE-DDDMNE-- 321

Qy        115 DTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKH 174
              ||| :||  |::|||||: || |: |:   : |  |  |:     |         | |  |
Db        322 DTVERIVRCIIQNEANAEALKEMLGDS---EGEGTVQLSS-----VDATSSLQDGAPSHH 373

Qy        175 VCGHHLHTVGGVVERDVCHRCRHKRWHFIKP-TNKSRESRPRRQGEVTVLSVGRFRVTKV 233
                  |||          |  | ||   ::   :|  :||| | || || ||||||||  :
Db        374 ------HTVHLGSAAPCLHCSRSKRPPLVRQGRSKEGKSRP-RTGETTVFSVGRFRVTHI 426

Qy        234 EHKSNQKERRSLMSV-----SGAETVNGEVPATPVKR 265
              | :    | ||           || |        :| :
Db        427 EKRYGLHEHRDGSPTDRSWGSGVSTTGSREYVSPADK 463


RESULT 6
US-09-205-258-634
; Sequence 634, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/09/205,258
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
```

```
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 634
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-205-258-634

  Query Match             10.3%;  Score 147;  DB 2;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-08;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        124 IMKNEANADVLKAMVADNSLYDPESPVTP 152
              |||||||||||||||||||||||||||||
Db          1 IMKNEANADVLKAMVADNSLYDPESPVTP 29


RESULT 7
US-10-004-860-634
; Sequence 634, Application US/10004860
; Patent No. 6914047
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/004,860
;  CURRENT FILING DATE:  2001-12-07
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 634
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-004-860-634

  Query Match             10.3%;  Score 147;  DB 2;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-08;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        124 IMKNEANADVLKAMVADNSLYDPESPVTP 152
              |||||||||||||||||||||||||||||
Db          1 IMKNEANADVLKAMVADNSLYDPESPVTP 29


RESULT 8
US-09-248-796A-15758
; Sequence 15758, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15758
;   LENGTH: 819
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-15758

  Query Match             7.3%;  Score 103.5;  DB 2;  Length 819;
  Best Local Similarity   27.1%;  Pred. No. 0.19;
  Matches   26;  Conservative   21;  Mismatches   38;  Indels   11;  Gaps    2;

Qy         91 EAEQDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPV 150
              : ::| :::| |: : ::  :|| ||| ::|       | : :  :  | |:  | ||:
Db        721 DKDEDKDKDKEEEDDDDEYNDENEDTVDELV-------AGFEKVVPLTEDKSMGTPSSPI 773

Qy        151 TPSTPGSP----PVSPGPLSPGGTPGKHVCGHHLHT 182
              |||| :|    | :    :|  |       |  | |
Db        774 IPSTPKTPKSITPTTSKSTTPTTTASTETIGKRLST 809


RESULT 9
US-10-104-047-2688
; Sequence 2688, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. 6943241el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2688
;   LENGTH: 447
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-104-047-2688

  Query Match             6.8%;  Score 97.5;  DB 2;  Length 447;
  Best Local Similarity   20.6%;  Pred. No. 0.32;
  Matches   72;  Conservative   37;  Mismatches   87;  Indels  153;  Gaps   16;

Qy         29 PDNGSSRTLH----------------------------SRTETTPSPSNDTGNGHPEYI 59
              ||| :|:|                            :| :::|  |  : :  |  |
Db         73 PDNTTSKTTDCLQTKGFSNSTEHKRGSVAQKVQEFNNCLNRGQSSPQRSYSSSHSSPAKI 132

Qy         60 AYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTV-G 118
                                    | | |    ||  :  ::  :  |  | : ||
Db        133 -------------------------QRATQEPVAKIEGTQESQMVGSSSTREKASTVLS 166

Qy        119 QIVHYIMKNEA---------------------NADVLKAM-----VADNSLY------- 144
              |||  |   ::                    :||| |:      |   ||:
Db        167 QIVASIQPPQSPPETPQSGPKACSVEELYAIPPDADVAKSTPKSTPVRPKSLFTSQPSGE 226

Qy        145 ---------------DPE-SPVTPS------TPGSPPVSPGPLSPGGTPGKHVCGHHLH 181
                           ||  |||||       :| | | :| |  |  |  |   :  |: |
Db        227 AEAPQTTDSPTTKVQKDPSIKPVTPSPSKLVTSPQSEPPAPFP-PPRSTSSPYHAGNLLQ 285

Qy        182 TVGGVVERDVCHRCRH-KRWHFIKPTNKSR----------------ESRPRRQGEVTVL 223
                          ||   |   |||: :|                :::|:|   ::
Db        286 -------------RHFTNW--TKPTSPTRSTEAESVLHSEGSRRAADAKPKRW--ISFK 327
```

```
Qy        224 SVGRFRVTKVE-HKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
Db        328 SFFRRRKTDEEDDKEKEREKGKLVGLDG--TVIHMLPPPPVQRHHWFTE 374


RESULT 10
US-08-985-950-6
; Sequence 6, Application US/08985950
; Patent No. 6140076
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  287 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-985-950-6

  Query Match             6.8%;  Score 97;  DB 2;  Length 287;
  Best Local Similarity   22.7%;  Pred. No. 0.19;
  Matches   58;  Conservative   45;  Mismatches  105;  Indels   48;  Gaps   14;

Qy         34 SRTLHSRTETTPSPSNDTGNGHPEYIAYALVPV-----FFIMGLFGVLICHLLKKKGYRC 88
Db         16 AQTIHTQEEDLPRPSISAEPG-------TVIPLGSHVTFVCRGPVGVQTFRLERES--RS 66

Qy         89 TTEAEQDIEEEKVEKIELN---DSVNENSDTVGQIVHY-IMKNEANADVLKAMVADNSLY 144
Db         67 TYNDTEDVSQASPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSEQSDYLELLVKETS-G 125

Qy        145 DPESPVT-PSTPGSPPVSP--------GPLSPGGTPGKHVCGHHLHTVGGV-VERDVCHR 194
Db        126 GPDSPDTEPGSSAGPTQRPSDNSHNEHAPASQG------LKAEHLYILIGVSVVFLFCLL 179

Qy        195 -----CRHKRWHFIK---PTNKSRESRPRRQGEVTV----LSVGRFRVTKVEHKSNQKER 242
Db        180 LLVLFCLH-RQNQIKQGPPRSKDEEQKPQQRPDLAVDVLERTADKATVNGLPEKDRETDT 238

Qy        243 RSLMSVSGAETVNGEV 258
Db        239 SALAAGSSQEVTYAQL 254


RESULT 11
US-08-985-950-8
; Sequence 8, Application US/08985950
; Patent No. 6140076
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  287 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-985-950-8

  Query Match             6.8%;  Score 97;  DB 2;  Length 287;
  Best Local Similarity   22.7%;  Pred. No. 0.19;
  Matches   58;  Conservative   45;  Mismatches  105;  Indels   48;  Gaps   14;

Qy         34 SRTLHSRTETTPSPSNDTGNGHPEYIAYALVPV-----FFIMGLFGVLICHLLKKKGYRC 88
Db         16 AQTIHTQEEDLPRPSISAEPG-------TVIPLGSHVTFVCRGPVGVQTFRLERES--RS 66

Qy         89 TTEAEQDIEEEKVEKIELN---DSVNENSDTVGQIVHY-IMKNEANADVLKAMVADNSLY 144
Db         67 TYNDTEDVSQASPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSEQSDYLELLVKETS-G 125

Qy        145 DPESPVT-PSTPGSPPVSP--------GPLSPGGTPGKHVCGHHLHTVGGV-VERDVCHR 194
Db        126 GPDSPDTEPGSSAGPTQRPSDNSHNEHAPASQG------LKAEHLYILIGVSVVFLFCLL 179

Qy        195 -----CRHKRWHFIK---PTNKSRESRPRRQGEVTV----LSVGRFRVTKVEHKSNQKER 242
Db        180 LLVLFCLH-RQNQIKQGPPRSKDEEQKPQQRPDLAVDVLERTADKATVNGLPEKDRETDT 238
```

```
Qy        243 RSLMSVSGAETVNGEV 258
              :| : |  |    ::
Db        239 SALAAGSSQEVTYAQL 254


RESULT 12
US-09-546-049-6
; Sequence 6, Application US/09546049
; Patent No. 6479638
;   GENERAL INFORMATION:
;        APPLICANT: Adema, Gosse Jan
;                   Meyaard, Linde
;                   Gorman, Daniel M.
;                   McClanahan, Terrill K.
;                   Zurawski, Sandra M.
;                   Zurawski, Gerard
;                   Lanier, Lewis L.
;                   Phillips Jr., Joseph H.
;        TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
;                            Related Reagents
;        NUMBER OF SEQUENCES: 22
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DNAX Research Institute
;             STREET: 901 California Avenue
;             CITY: Palo Alto
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94304-1104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/546,049
;             FILING DATE: 10-Apr-2000
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/985,950
;             FILING DATE: 05-DEC-1997
;             APPLICATION NUMBER: US 60/041,279
;             FILING DATE: 21-MARCH-1997
;             APPLICATION NUMBER: US 60/033,181
;             FILING DATE: 16-DEC-1996
;             APPLICATION NUMBER: US 60/032,252
;             FILING DATE: 06-DEC-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ching, Edwin P.
;             REGISTRATION NUMBER: 34,090
;             REFERENCE/DOCKET NUMBER: DX0670K
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650)852-9196
;             TELEFAX: (650)496-1204
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 287 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-546-049-6

  Query Match             6.8%;  Score 97;  DB 2;  Length 287;
  Best Local Similarity   22.7%;  Pred. No. 0.19;
  Matches   58;  Conservative   45;  Mismatches  105;  Indels   48;  Gaps   14;

Qy         34 SRTLHSRTETTPSPSNDTGNGHPEYIAYALVPV-----FFIMGLFGVLICHLLKKKGYRC 88
              ::|:|:: |  | ||     |        ::|:     |   |  ||    | ::  |
Db         16 AQTIHTQEEDLPRPSISAEPG-------TVIPLGSHVTFVCRGPVGVQTFRLERES--RS 66

Qy         89 TTEAEQDIEEEKVEKIELN---DSVNENSDTVGQIVHY-IMKNEANADVLKAMVADNSLY 144
              |    :|: :    : |     |||:| :    : ::|   |    :| |: :| : |
Db         67 TYNDTEDVSQASPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSEQSDYLELLVKETS-G 125

Qy        145 DPESPVT-PSTPGSPPVSP--------GPLSPGGTPGKHVCGHHLHTVGGV-VERDVCHR 194
              |:|| | |  :    |  |        | | |      :   ||: : || |   |
Db        126 GPDSPDTEPGSSAGPTQRPSDNSHNEHAPASQG------LKAEHLYILIGVSVVFLFCLL 179

Qy        195 -----CRHKRWHFIK---PTNKSRESRPRRQGEVTV----LSVGRFRVTKVEHKSNQKER 242
                  | | | : ||   | :|  | :|::: :: |     :  :  |  :  |  : :
Db        180 LLVLFCLH-RQNQIKQGPPRSKDEEQKPQQRPDLAVDVLERTADKATVNGLPEKDRETDT 238

Qy        243 RSLMSVSGAETVNGEV 258
              :| : |  |    ::
Db        239 SALAAGSSQEVTYAQL 254


RESULT 13
US-09-546-049-8
; Sequence 8, Application US/09546049
; Patent No. 6479638
;   GENERAL INFORMATION:
;        APPLICANT: Adema, Gosse Jan
;                   Meyaard, Linde
;                   Gorman, Daniel M.
;                   McClanahan, Terrill K.
;                   Zurawski, Sandra M.
;                   Zurawski, Gerard
;                   Lanier, Lewis L.
;                   Phillips Jr., Joseph H.
;        TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
;                            Related Reagents
;        NUMBER OF SEQUENCES: 22
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DNAX Research Institute
;             STREET: 901 California Avenue
;             CITY: Palo Alto
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94304-1104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/546,049
;             FILING DATE: 10-Apr-2000
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/985,950
;             FILING DATE: 05-DEC-1997
;             APPLICATION NUMBER: US 60/041,279
;             FILING DATE: 21-MARCH-1997
;             APPLICATION NUMBER: US 60/033,181
;             FILING DATE: 16-DEC-1996
;             APPLICATION NUMBER: US 60/032,252
;             FILING DATE: 06-DEC-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ching, Edwin P.
;             REGISTRATION NUMBER: 34,090
;             REFERENCE/DOCKET NUMBER: DX0670K
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650)852-9196
;             TELEFAX: (650)496-1204
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 287 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-546-049-8

  Query Match             6.8%;  Score 97;  DB 2;  Length 287;
  Best Local Similarity   22.7%;  Pred. No. 0.19;
  Matches   58;  Conservative   45;  Mismatches  105;  Indels   48;  Gaps   14;
```

```
Qy          34 SRTLHSRTETTPSPSNDTGNGHPEYIAYALVPV-----FFIMGLFGVLICHLLKKKGYRC 88
               ::|:|:: |  | ||     |        ::|:     |   |  ||    | ::   |
Db          16 AQTIHTQEEDLPRPSISAEPG-------TVIPLGSHVTFVCRGPVGVQTFRLERES--RS 66

Qy          89 TTEAEQDIEEEKVEKIELN---DSVNENSDTVGQIVHY-IMKNEANADVLKAMVADNSLY 144
               |    :|: :    : |     |||:| :    : ::|    |    :| |: :| : |
Db          67 TYNDTEDVSQASPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSEQSDYLELLVKETS-G 125

Qy         145 DPESPVT-PSTPGSPPVSP--------GPLSPGGTPGKHVCGHHLHTVGGV-VERDVCHR 194
               |:|| | | :   |   |          | | |      :   ||: : || | |
Db         126 GPDSPDTEPGSSAGPTQRPSDNSHNEHAPASQG------LKAEHLYILIGVSVVFLFCLL 179

Qy         195 -----CRHKRWHFIK---PTNKSRESRPRRQGEVTV----LSVGRFRVTKVEHKSNQKER 242
                    | | | : ||   | :|  | :|::: :: |     :  : |  : |  : :
Db         180 LLVLFCLH-RQNQIKQGPPRSKDEEQKPQQRPDLAVDVLERTADKATVNGLPEKDRETDT 238

Qy         243 RSLMSVSGAETVNGEV 258
               :| : |  |    ::
Db         239 SALAAGSSQEVTYAQL 254


RESULT 14
US-08-890-865A-1
; Sequence 1, Application US/08890865A
; Patent No. 6307019
;  GENERAL INFORMATION:
;    APPLICANT:  Constantini, Franklin
;    APPLICANT:  Zeng, Li
;    TITLE OF INVENTION:  AXIN GENE AND USES THEREOF
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  US
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,865A
;      FILING DATE:  10-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/54249
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)278-0400
;      TELEFAX:  (212)391-0526
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  992 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Protein
US-08-890-865A-1

  Query Match             6.7%;  Score 95;  DB 2;  Length 992;
  Best Local Similarity   19.6%;  Pred. No. 1.8;
  Matches   45;  Conservative   35;  Mismatches   94;  Indels   56;  Gaps    8;

Qy          89 TTEAEQDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKN-----------EANADVLKA 136
               | |||: : ||:::: : :    :     | :  : : :             :     |
Db         531 TREAEEKL-EERLKRVRMEEEGEDGEMPSGPMASHKLPSVPAWHHFPPRYVDMGCSGLRD 589

Qy         137 MVADN--SLYDPESPVTPSTPGSPPVSPGPLSP-----------GGTPGKHVCGHHLHTV 183
                 :|  |: |          |||      ||  ||          |||   |  | |: :
Db         590 AHEENPESILDEHVQRVMRTPGCQSPGPGHRSPDSGHVAKTAVLGGTASGH--GKHVPKL 647

Qy         184 GGVVERDVCHRCRHKRWHFIKPTNKSRE----------------------SRPRRQGE- 219
               |  ::    |  ||   |    : : :|                     ::||   |
Db         648 GLKLDTAGLHHHRHVHHHVHHNSARPKEQMEAEVARRVQSSFSWGPETHGHAKPRSYSEN 707

Qy         220 -VTVLSVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEVPATPVKRERS 268
                | || |          |::   :|:       :  | |||:|    |::
Db         708 AGTTLSAGDL---PFGGKTSAPSKRNTKKAESGKNANAEVPSTTEDAEKN 754


RESULT 15
US-08-015-985-3
; Sequence 3, Application US/08015985
; Patent No. 5538886
;  GENERAL INFORMATION:
;    APPLICANT:  Schlessinger, Joseph
;    APPLICANT:  Sap, Jan M.
;    TITLE OF INVENTION:  NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
;    TITLE OF INVENTION:  PHOSPHATASE-ALPHA
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 AVENUE OF THE AMERICAS
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/015,985
;      FILING DATE:  10-FEB-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7683-020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-015-985-3

  Query Match             6.6%;  Score 94;  DB 1;  Length 793;
  Best Local Similarity   19.0%;  Pred. No. 1.6;
  Matches   63;  Conservative   51;  Mismatches  124;  Indels   94;  Gaps   13;

Qy          13 AAAVFVGGAVSSP------------LVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYI 59
               ||:   ||   ||             :  |  |:| |  :  || | |:::|       |
Db          91 AASTESGGTTISPNGSWLIENQFTDAITEPWEGNSSTAATTPETFP-PADET-----PII 144

Qy          60 AYALVPVFFIMGLFGVLICHLLKKKGY----------RCTTEAEQDIEEEKVEKIELNDS 109
               |  :     :: :| ::: ::|: | |           | :    :|:| : |  :  : |
Db         145 AVMVALSSLLVIVFIIIVLYMLRFKKYKQAGSHSNSFRLSNGRTEDVEPQSVPLLARSPS 204

Qy         110 VNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGG 169
               |            :  ::    :: : |  || |:  |     |: |
Db         205 TNRKYPP-------LPVDKLEEEINRRMADDNKLFREEFNALPACPIQATCEAASKEENK 257

Qy         170 TPGKHV----CGH---HLHTVGGVVERDVCHRC-----RHKR----------------WH 201
                  ::|      |   ||  | || : |  :       : |                  |
Db         258 EKNRYVNILPYDHSRVHLTPVEGVPDSDYINASFINGYQEKNKFIAAQGPKEETVNDFWR 317

Qy         202 FIKPTNKS--------RESRPRR-------QG-----------EVTVL---SVGRFRVT 231
```

```
                |   | :       :| :  :       ||           :||||   :| :| :
Db        318 MIWEQNTATIVMVTNLKERKECKCAQYWPDQGCWTYGNVRVSVEDVTVLVDYTVRKFSIQ 377

Qy        232 KVEHKSNQKERRSLMSVSGAETVNGEVPATPV 263
              :|   :|:| :| :         :  || ||:
Db        378 QVGDVTNRKPQRLITQFHFTSWPDFGVPFTPI 409


Search completed: February 23, 2006, 12:09:26
Job time : 48 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 23, 2006, 12:00:48 ; Search time 231 Seconds
(without alignments)
827.698 Million cell updates/sec

Title: US-10-620-562-2
Perfect score: 1426
Sequence: 1 MAPRALPGSAVLAAAVFVGG.....ETVNGEVPATPVKRERSGTE 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1426	100.0	271	2	Q8IUW5_HUMAN	Q8iuw5 homo sapien
2	1421	99.6	271	2	Q8NBK1_HUMAN	Q8nbk1 homo sapien
3	986	69.1	202	2	Q8K2J7_MOUSE	Q8k2j7 mus musculu
4	854.5	59.9	288	2	Q5F3A4_CHICK	Q5f3a4 gallus gall
5	699	49.0	143	2	Q8R3S5_MOUSE	Q8r3s5 mus musculu
6	669	46.9	223	2	Q7T0N2_XENLA	Q7t0n2 xenopus lae
7	641.5	45.0	224	2	Q6NRQ3_XENLA	Q6nrq3 xenopus lae
8	581	40.7	280	2	Q5M7X6_BRARE	Q5m7x6 brachydanio
9	422.5	29.6	281	2	Q4SRN7_TETNG	Q4srn7 tetraodon n
10	309	21.7	430	1	TR19L_MACFA	Q9n092 macaca fasc
11	307	21.5	430	1	TR19L_HUMAN	Q969z4 homo sapien
12	295.5	20.7	436	1	TR19L_MOUSE	Q8bx43 mus musculu
13	289	20.3	294	2	Q8CFT0_MOUSE	Q8cft0 mus musculu
14	289	20.3	303	2	Q8BRJ3_MOUSE	Q8brj3 mus musculu
15	289	20.3	325	2	Q6DIW4_XENTR	Q6diw4 xenopus tro
16	286	20.1	302	2	Q5FVJ4_RAT	Q5fvj4 rattus norv
17	286	20.1	327	2	Q68EV2_XENLA	Q68ev2 xenopus lae
18	284	19.9	303	2	Q6P4E7_HUMAN	Q6p4e7 homo sapien
19	284	19.9	303	2	Q6UXY2_HUMAN	Q6uxy2 homo sapien
20	284	19.9	303	2	Q8NC24_HUMAN	Q8nc24 homo sapien
21	282.5	19.8	401	2	Q6NZ11_BRARE	Q6nz11 brachydanio
22	158.5	11.1	163	2	Q4SEL6_TETNG	Q4sel6 tetraodon n
23	152.5	10.7	52	2	Q4SZ05_TETNG	Q4sz05 tetraodon n
24	138	9.7	29	2	Q4W5D7_HUMAN	Q4w5d7 homo sapien
25	110.5	7.7	2287	2	Q8I520_PLAF7	Q8i520 plasmodium
26	110	7.7	1244	2	Q51RT1_MAGGR	Q51rt1 magnaporthe
27	109	7.6	392	2	Q9I8G8_CHEMY	Q9i8g8 chelonia my
28	109	7.6	875	2	Q5FWN0_XENLA	Q5fwn0 xenopus lae
29	108	7.6	800	2	Q5AH61_CANAL	Q5ah61 candida alb
30	108	7.6	800	2	Q5AIP1_CANAL	Q5aip1 candida alb
31	107.5	7.5	515	2	Q4H3L1_CIOIN	Q4h3l1 ciona intes
32	106.5	7.5	805	2	Q55MF2_CRYNE	Q55mf2 cryptococcu
33	106	7.4	795	2	Q6EUQ1_ORYSA	Q6euq1 oryza sativ
34	105.5	7.4	805	2	Q5K8C2_CRYNE	Q5k8c2 cryptococcu
35	105.5	7.4	1083	2	Q12075_PNECA	Q12075 pneumocysti
36	105.5	7.4	3140	2	Q52V19_9POTV	Q52v19 plum pox vi
37	104.5	7.3	1679	2	Q26607_SCHMA	Q26607 schistosoma
38	104.5	7.3	1791	2	Q9U6D4_PLABE	Q9u6d4 plasmodium
39	104	7.3	1787	2	Q25645_PLABE	Q25645 plasmodium
40	103.5	7.3	384	2	Q8V9I8_9POTV	Q8v9i8 plum pox vi
41	103	7.2	749	2	Q72SP6_LEPIC	Q72sp6 leptospira
42	103	7.2	749	2	Q8F2V5_LEPIN	Q8f2v5 leptospira
43	103	7.2	1279	2	Q4S7I1_TETNG	Q4s7i1 tetraodon n
44	102.5	7.2	751	2	Q5E9H4_BOVIN	Q5e9h4 bos taurus
45	102.5	7.2	1829	2	Q7XE96_ORYSA	Q7xe96 oryza sativ

ALIGNMENTS

RESULT 1
Q8IUW5_HUMAN
ID Q8IUW5_HUMAN PRELIMINARY; PRT; 271 AA.
AC Q8IUW5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to expressed sequence AA536743.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC039540; AAH39540.1; -; mRNA.
DR Ensembl; ENSG00000181826; Homo sapiens.
SQ SEQUENCE 271 AA; 29340 MW; D85AAA0E86349EA7 CRC64;

Query Match 100.0%; Score 1426; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 3.2e-100;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
Db 1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60

```
Qy          61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy         121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180

Qy         181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240

Qy         241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
               |||||||||||||||||||||||||||||||
Db         241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271


RESULT 2
Q8NBK1_HUMAN
ID   Q8NBK1_HUMAN PRELIMINARY;      PRT;   271 AA.
AC   Q8NBK1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   CDNA PSEC0162 fis, clone PLACE1009772.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Placenta;
RA   Otsuki T., Ota T., Nishikawa T., Hayashi K., Suzuki Y., Yamamoto J.,
RA   Wakamatsu A., Kimura K., Sakamoto K., Hatano N., Kawai Y., Ishii S.,
RA   Saito K., Kojima S., Sugiyama T., Ono T., Okano K., Yoshikawa Y.,
RA   Aotsuka S., Sasaki N., Hattori A., Okumura K., Nagai K., Sugano S.,
RA   Isogai T.;
RT   "Signal Sequence and Keyword Trap in silico for Selection of Full-
RT   Length Human cDNAs Encoding Secretion or Membrane Proteins from Oligo-
RT   Capped cDNA Libraries.";
RL   DNA Res. 12:117-126(2005).
DR   EMBL; AK075468; BAC11638.1; -; mRNA.
SQ   SEQUENCE   271 AA;  29310 MW;  385AAA0A97258FBD CRC64;

  Query Match             99.6%;  Score 1421;  DB 2;  Length 271;
  Best Local Similarity   99.6%;  Pred. No. 7.7e-100;
  Matches  270;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSNDTGNGHPEYIA 60
               |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db           1 MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRALHSRTETTPSPSNDTGNGHPEYIA 60

Qy          61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 YALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQI 120

Qy         121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 VHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHL 180

Qy         181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 HTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240

Qy         241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
               |||||||||||||||||||||||||||||||
Db         241 ERRSLMSVSGAETVNGEVPATPVKRERSGTE 271


RESULT 3
Q8K2J7_MOUSE
ID   Q8K2J7_MOUSE PRELIMINARY;      PRT;   202 AA.
AC   Q8K2J7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Expressed sequence AA536743.
GN   Name=AA536743;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FVB/N-3, and C57BL/6J;
RC   TISSUE=Embryo, and
RC   Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by
RC   biopsy.;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FVB/N-3;
RC   TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by
RC   biopsy.;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RA   Director MGC Project;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC031198; AAH31198.1; -; mRNA.
DR   EMBL; BC066137; AAH66137.1; -; mRNA.
DR   EMBL; BC066160; AAH66160.1; -; mRNA.
DR   Ensembl; ENSMUSG00000047881; Mus musculus.
DR   MGI; MGI:2140767; AA536743.
SQ   SEQUENCE   202 AA;  22263 MW;  092F50EAF995E6B8 CRC64;

  Query Match             69.1%;  Score 986;  DB 2;  Length 202;
  Best Local Similarity   91.1%;  Pred. No. 6.3e-67;
  Matches  184;  Conservative    9;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          70 MGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEA 129
               ||| |||||||||||||||||::||||||||||||:|||||||||| |||||||
Db           1 MGLLGVLICHLLKKKGYRCTTEAEQEVEEEKVEKIELNDSINENSDTVGQIVQYIMKNEA 60

Qy         130 NADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHLHTVGGVVER 189
               |||:|||||||||: | |||||||||||||||||||| ||||||||||||||||||||
Db          61 NADILKAMVADNSVGDIESPVTPSTPGSPPVSPGPLSPGATPGKHVCGHHLHTVGGVVER 120

Qy         190 DVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLMSVS 249
               ||| ||||||||||||::| ||||||||||||||||||||||||||||||||||||||
Db         121 DVCQRCRHKRWHFIKPTNKTKEGRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLMSVS 180
```

```
Qy        250 GAETVNGEVPATPVKRERSGTE 271
              | |:|||:|||||||||||| ||
Db        181 GIESVNGDVPATPVKRERSDTE 202


RESULT 4
Q5F3A4_CHICK
ID   Q5F3A4_CHICK PRELIMINARY;      PRT;   288 AA.
AC   Q5F3A4;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=RCJMB04_24l24;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CB; TISSUE=Bursa;
RA   Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA   Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA   Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT   "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT   genefunction analysis.";
RL   Genome Biol. 6:R6-R6(2005).
DR   EMBL; AJ851746; CAH65380.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   288 AA;  31157 MW;  07DC2D83E938C7BD CRC64;

  Query Match           59.9%;  Score 854.5;  DB 2;  Length 288;
  Best Local Similarity   61.0%;  Pred. No. 9.3e-57;
  Matches  177;  Conservative   36;  Mismatches   50;  Indels   27;  Gaps    8;

Qy          1 MAPRALPGSAVLA-----AAVFVGGAVSSPLVAPDNGSSRTLHSR--TETTPSPSNDTGN 53
              ||| |  |   :|      ||::|          | |  : | ||    || :  |:
Db          1 MAPPAASGIPSIAPSLGPTAVWLGN-------RSDLGDVQALASRDLPTTTVTAGNNNKP 53

Qy         54 GHPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKV-EKIELNDSVNE 112
                | ||:|: |||||||||| |:||||:|||||||||||||:  ||||  ||||:|:: :|
Db         54 EHLEYVAPVLVPVFFIMGLLGILICHVLKKKGYRCTTEAEEVEEEEKPDEKIEMNETAHE 113

Qy        113 NSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPG-PLSPGGTP 171
              ||||||||::|||||||||||||||||||:|:::||||::|: ||| | ||| |||||
Db        114 NSDTVGQIINYIMKNEANADVLKAMVADSSVFEPESPMSPNAPGS-PTSPGSPLSPGAAS 172

Qy        172 GKHVC-GHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRV 230
                || | |||||||||||:|:||| || ||||| ||| :||:| |  | ||||||||||||
Db        173 LKHNCKGHHLHTVGGVIEKDVCSRCSHKRWHHIKPAHKSKEHRRSRLGEVTVLSVGRFRV 232

Qy        231 TKVEHKSNQKERRSLMSVSGAETVNGEVPATPVKRE---------RSGTE 271
              |||||||| |||:|||||:| | :||::||||||:|         : ||:
Db        233 TKVEHKSNSKERKSLMSVTGVEGLNGDMPATPVKQEAKEAPATPVKEGTQ 282


RESULT 5
Q8R3S5_MOUSE
ID   Q8R3S5_MOUSE PRELIMINARY;      PRT;   143 AA.
AC   Q8R3S5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   AA536743 protein (Fragment).
GN   Name=AA536743;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CZECH II;
RC   TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CZECH II;
RC   TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024679; AAH24679.1; -; mRNA.
DR   MGI; MGI:2140767; AA536743.
FT   NON_TER       1       1
SQ   SEQUENCE   143 AA;  15527 MW;  59547DA0F91D5D42 CRC64;

  Query Match           49.0%;  Score 699;  DB 2;  Length 143;
  Best Local Similarity   90.9%;  Pred. No. 2.8e-45;
  Matches  130;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        129 ANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHLHTVGGVVE 188
              ||||:|||||||||: | ||||||||||||||||||||| ||||||||||||||||||
Db          1 ANADILKAMVADNSVGDIESPVTPSTPGSPPVSPGPLSPGATPGKHVCGHHLHTVGGVVE 60

Qy        189 RDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLMSV 248
              |||| ||||||||||||||::| ||||||||||||||||||||||||||||||||||||
Db         61 RDVCQRCRHKRWHFIKPTNKTKEGRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLMSV 120

Qy        249 SGAETVNGEVPATPVKRERSGTE 271
              || |:|||:|||||||||||| ||
Db        121 SGIESVNGDVPATPVKRERSDTE 143


RESULT 6
Q7T0N2_XENLA
ID   Q7T0N2_XENLA PRELIMINARY;      PRT;   223 AA.
AC   Q7T0N2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MGC69140 protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
```

```
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Ovary;
RA   Klein S., Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC056115; AAH56115.1; -; mRNA.
SQ   SEQUENCE   223 AA;  24273 MW;  5EB8CD32032367AB CRC64;

  Query Match             46.9%;  Score 669;  DB 2;  Length 223;
  Best Local Similarity   65.4%;  Pred. No. 8.7e-43;
  Matches  142;  Conservative   21;  Mismatches   36;  Indels   18;  Gaps    9;

Qy         70 MGLFGVLICHLLKKKGYRCTTEAEQDIEEEKV-EKIELNDSVNE-NSDTVGQIVHYIMKN 127
              ||| |||||||||||||||||||||  :||: | ||||: :|  : |:|||||| :||||
Db          1 MGLVGVLICHLLKKKGYRCTTEAEPPVEEKIVGEKIEMQESTGDTNTDTVGQIVDFIMKN 60

Qy        128 EANADVLKAMVADNSL------YDPESPVTPSTPGSP--PVSP-GPLSPGGTPGKHVC-G 177
              |||||:|||||||||:      :|||| ||:|||||  | :|  | ||  || || | |
Db         61 EANADILKAMVADNSIVGDTSVFDPESPTTPNTPGSPITPDTPISPTSPAETPSKHSCRG 120

Qy        178 HHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKS 237
              ||||||||| ||: | || :|||||:|   | :| |   || ||||||||||||||| ||
Db        121 HHLHTVGGVAERNACTRCTNKRWHFLKSPQKHKEPRRSHQGAVTVLSVGRFRVTKVEPKS 180

Qy        238 NQKERRSLMSVSGAETVNGEVPATPVK---RERSGTE 271
                 |||: ||:   :|  |||||||||    |:||||:
Db        181 --KERKRLMA-DRSEGTNGEVPATPVSVDTRQRSGTD 214


RESULT 7
Q6NRQ3_XENLA
ID   Q6NRQ3_XENLA PRELIMINARY;      PRT;   224 AA.
AC   Q6NRQ3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC83066 protein.
GN   Name=MGC83066;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Ovary;
RA   Klein S., Strausberg R.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC070680; AAH70680.1; -; mRNA.
SQ   SEQUENCE   224 AA;  24512 MW;  516BF3010B11D9F0 CRC64;

  Query Match             45.0%;  Score 641.5;  DB 2;  Length 224;
  Best Local Similarity   63.8%;  Pred. No. 1.1e-40;
  Matches  139;  Conservative   21;  Mismatches   39;  Indels   19;  Gaps    9;

Qy         70 MGLFGVLICHLLKKKGYRCTTEAEQDIEEEKV--EKIELNDSVNE-NSDTVGQIVHYIMK 126
              ||| |||||||||||||||||||||    |||:  ||||:::|  : |:|||| ||::|||
Db          1 MGLVGVLICHLLKKKGYRCTTEAEPPPVEEKIVGEKIEMHESNGDTNTDTVGHIVNFIMK 60

Qy        127 NEANADVLKAMVADN------SLYDPESPVTPSTPGS--PPVSP-GPLSPGGTPGKHVC- 176
              ||||||||||||||      |::|||| ||:|| |  || :|  | ||  || || |
Db         61 NEANADVLKAMVADNSVVGDTSVFDPESPTTPNTPTSPLPPDTPISPTSPTETPSKHSCR 120

Qy        177 GHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHK 236
              | |||||||| ||:|| || :|||| ||   | :| |   || ||||||||||||||| |
Db        121 GSHLHTVGGVAERNVCSRCTNKRWHLIKSPQKHKEPRRSHQGAVTVLSVGRFRVTKVEPK 180

Qy        237 SNQKERRSLMSVSGAETVNGEVPATPVK---RERSGTE 271
              |  |||: ||:    |  |||:| |||:   |:||||:
Db        181 S--KERKRLMA-DRTEATNGEIPVTPVRVDTRQRSGTD 215


RESULT 8
Q5M7X6_BRARE
ID   Q5M7X6_BRARE PRELIMINARY;      PRT;   280 AA.
AC   Q5M7X6;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   LOC553285 protein (Fragment).
GN   Name=LOC553285;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Whole;
```

```
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Whole;
RG   NIH MGC Project;
RL   Submitted (DEC-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC088389; AAH88389.1; -; mRNA.
FT   NON_TER       1       1
SQ   SEQUENCE   280 AA;  30423 MW;  251CCA8028435E76 CRC64;

  Query Match             40.7%;  Score 581;  DB 2;  Length 280;
  Best Local Similarity   49.6%;  Pred. No. 5.5e-36;
  Matches  123;  Conservative   43;  Mismatches   62;  Indels   20;  Gaps    9;

Qy         29 PDNGSSRTLHSRTETTPSPSNDTGNGHPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRC 88
              |  ||| : :|   ||   :  :|:|||| ||: ||| ||:||| |||||||:||:|||||
Db         22 PMAGSSLSTNS---TTRGGNGHSGDGHPENIAFILVPFFFLMGLLGVLICHILKRKGYRC 78

Qy         89 TTEAEQDIE------EEKVEKIELNDSVNE-NSDTVGQIVHYIMKNEANADVLKAMVADN 141
              |||||:: :      :|::|| :|||: :| |:|||||||||||||||:| ||||| |
Db         79 TTEAEEEEQLEKERRDEELEKGDLNDTFSEGNADTVGQIVHYIMKNEANSDALKAMVQD- 137

Qy        142 SLYDPESPVTPSTPGSPPVSPGPLSPG--GTPGKHVCGHHLHTVGGVV-ERDVCHRCRHK 198
              |:     ||||:|| :|    | :||   |  || | :||||:||:   :::|||| |
Db        138 SIDSEGGPVTPTTPNTPTSPESPAAPGLPPTAAKHTC-NHLHTIGGISGHKNICHRCNQK 196

Qy        199 RWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEV 258
              :|  :: ::  :  |    |||||||||||||| : |:  :|||:|:        || |
Db        197 KWPLLRRSSSKKLDRRSHVGEVTVLSVGRFRVTKCDPKT-ARERRTLLITE----PNGSV 251

Qy        259 PATPVKRE 266
                :|   :
Db        252 SPSPASTQ 259


RESULT 9
Q4SRN7_TETNG
ID   Q4SRN7_TETNG PRELIMINARY;      PRT;   281 AA.
AC   Q4SRN7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome undetermined SCAF14506, whole genome shotgun sequence.
GN   ORFNames=GSTENG00013829001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014506; CAF96695.1; -; Genomic_DNA.
SQ   SEQUENCE   281 AA;  30111 MW;  7B09ACCF514A0325 CRC64;

  Query Match             29.6%;  Score 422.5;  DB 2;  Length 281;
  Best Local Similarity   46.5%;  Pred. No. 6.2e-24;
  Matches   93;  Conservative   33;  Mismatches   53;  Indels   21;  Gaps    8;

Qy         41 TETTPSPSNDTGNG------HPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTE--- 91
              |  :   : |  ||       | |:|: ||||||::|| ||:|||:||:||||||||
Db          5 TAVSQHNATDGKNGGGADGTQPNYVAFVLVPVFFLLGLLGVVICHVLKRKGYRCTTEPQD 64

Qy         92 -AEQDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPV 150
               ::  || | |::   |: : : ||:||||  || ||||:| ||||| :||:     |:
Db         65 GGKEAWEEGKDEEMG-EDADDSDHDTLGQIVRCIMSNEANSDALKAMVHENSVDSDGPPL 123

Qy        151 TPSTPGSPPVSPGPLSPGGTPG--KHVCGHHLHTVGGVVERDVCHRCRHKRWHFI-KPTN 207
              ||    |||::  |:||   ||  || | |||||||   :: | ||  |:|  : :|:
Db        124 TP----SPPMT--PVSPDAPPGAAKHTCS-HLHTVGGTGLKNTCTRCSQKKWPLMRRPSP 176

Qy        208 KSRESRPRRQGEVTVLSVGR 227
              :  | : |  ||||||:|||
Db        177 RKAEQQRRSAGEVTVLAVGR 196


RESULT 10
TR19L_MACFA
ID   TR19L_MACFA    STANDARD;      PRT;   430 AA.
AC   Q9N092;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Tumor necrosis factor receptor superfamily member 19L precursor
DE   (Receptor expressed in lymphoid tissues).
GN   Name=TNFRSF19L; Synonyms=RELT; ORFNames=QccE-10051;
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Brain cortex;
RX   MEDLINE=21458551; PubMed=11574149; DOI=10.1016/S0378-1119(01)00665-5;
RA   Osada N., Hida M., Kususda J., Tanuma R., Iseki K., Hirata M.,
RA   Suto Y., Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K.;
RT   "Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
RT   chromosomes.";
RL   Gene 275:31-37(2001).
RN   [2]
RP   ERRATUM.
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirata M., Suto Y.,
RA   Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K., Kususda J.;
```

```
RL   Gene 278:267-267(2001).
CC   -!- FUNCTION: Mediates activation of NF-kappa-B (By similarity). May
CC       play a role in T-cell activation.
CC   -!- SUBUNIT: Associates with TRAF1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC   -!- SIMILARITY: Contains 1 TNFR-Cys repeat.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB046039; BAB01621.1; -; mRNA.
DR   HSSP; Q92956; 1JMA.
DR   InterPro; IPR001368; TNFR_c6.
DR   SMART; SM00208; TNFR; 1.
DR   PROSITE; PS00652; TNFR_NGFR_1; FALSE_NEG.
DR   PROSITE; PS50050; TNFR_NGFR_2; FALSE_NEG.
KW   Glycoprotein; Receptor; Signal; Transmembrane.
FT   SIGNAL        1     26       By similarity.
FT   CHAIN        27    430       Tumor necrosis factor receptor
FT                                superfamily member 19L.
FT   TOPO_DOM     27    162       Extracellular (Potential).
FT   TRANSMEM    163    183       Potential.
FT   TOPO_DOM    184    430       Cytoplasmic (Potential).
FT   REPEAT       50     90       TNFR-Cys.
FT   CARBOHYD    149    149       N-linked (GlcNAc. . .) (Potential).
FT   DISULFID     51     65       By similarity.
FT   DISULFID     71     90       By similarity.
SQ   SEQUENCE   430 AA;  45851 MW;  BA8DE92593E1E859 CRC64;

  Query Match             21.7%;  Score 309;  DB 1;  Length 430;
  Best Local Similarity   31.0%;  Pred. No. 4.2e-15;
  Matches   84;  Conservative   37;  Mismatches   68;  Indels   82;  Gaps   12;

Qy         44 TPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
              |  | | |  | ||     |   |:|||| :||| |:|:|:|||:||| ||   |
Db        144 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 203

Qy         94 ------------QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
                        :|:                | ||:| :|  | : : ||  |: :: :
Db        204 GGGGSGINPAYRTEDV---------------NEDTIGVLVRLITEKKENAAALEELLKE 247

Qy        141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCR 196
                 |  :  |  |     |||  | :|   |  |:|   ||||||  |:  :    | ||
Db        248 ---YHSKQLVQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCS 299

Qy        197 HKRWHFI-----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKS 237
               |:|  :                   ||   :  ::  ||||:|:||||||| ::
Db        300 QKKWPEVLLSPEAVAATTSAPSFLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI---- 355

Qy        238 NQKERRSLMSVSGAETVN------GEVPATP 262
                |:|:   ||  :|:       |::| :|
Db        356 --PEQRTGSMVSEVKTITEAGPSAGDLPDSP 384


RESULT 11
TR19L_HUMAN
ID   TR19L_HUMAN    STANDARD;      PRT;   430 AA.
AC   Q969Z4; Q86V34; Q96JU1; Q9BUX7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Tumor necrosis factor receptor superfamily member 19L precursor
DE   (Receptor expressed in lymphoid tissues).
GN   Name=TNFRSF19L; Synonyms=RELT;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, PROTEIN SEQUENCE OF N-TERMINUS, AND INTERACTION
RP   WITH TRAF1.
RC   TISSUE=Lymphoma;
RX   MEDLINE=21213541; PubMed=11313261; DOI=10.1182/blood.V97.9.2702;
RA   Sica G.L., Zhu G., Tamada K., Liu D., Ni J., Chen L.;
RT   "RELT, a new member of the tumor necrosis factor receptor superfamily,
RT   is selectively expressed in hematopoietic tissues and activates
RT   transcription factor NF-kappaB.";
RL   Blood 97:2702-2707(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Retinoblastoma;
RX   PubMed=14702039; DOI=10.1038/ng1285;
RA   Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA   Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA   Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA   Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA   Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA   Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA   Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA   Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA   Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA   Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA   Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA   Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA   Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA   Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA   Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA   Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA   Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA   Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA   Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA   Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA   Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA   Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA   Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA   Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA   Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA   Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT   "Complete sequencing and characterization of 21,243 full-length human
RT   cDNAs.";
RL   Nat. Genet. 36:40-45(2004).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANT VAL-332.
RC   TISSUE=Blood, Colon, and Eye;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 121-430.
RC   TISSUE=Spleen;
RA   Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT   "The nucleotide sequence of a long cDNA clone isolated from human
```

```
RT   spleen.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   PROTEIN SEQUENCE OF 26-40.
RX   PubMed=15340161; DOI=10.1110/ps.04682504;
RA   Zhang Z., Henzel W.J.;
RT   "Signal peptide prediction based on analysis of experimentally
RT   verified cleavage sites.";
RL   Protein Sci. 13:2819-2824(2004).
CC   -!- FUNCTION: Mediates activation of NF-kappa-B. May play a role in T-
CC       cell activation.
CC   -!- SUBUNIT: Associates with TRAF1.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: Highest levels are in spleen, lymph node,
CC       thymus, peripheral blood leukocytes, bone marrow and fetal liver.
CC       Very low levels in skeletal muscle, testis and colon. Not detected
CC       in brain, kidney and pancreas.
CC   -!- SIMILARITY: Contains 1 TNFR-Cys repeat.
CC   -!- CAUTION: Ref.4 sequence differs from that shown due to several
CC       frameshifts.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF319553; AAK77356.1; -; mRNA.
DR   EMBL; AK027899; BAB55441.1; -; mRNA.
DR   EMBL; BC001812; -; NOT_ANNOTATED_CDS; mRNA.
DR   EMBL; BC017279; AAH17279.1; -; mRNA.
DR   EMBL; BC051810; AAH51810.2; -; mRNA.
DR   EMBL; AK074128; BAB84954.1; ALT_FRAME; mRNA.
DR   HSSP; Q92956; 1JMA.
DR   Ensembl; ENSG00000054967; Homo sapiens.
DR   HGNC; HGNC:13764; TNFRSF19L.
DR   H-InvDB; HIX0009918; -.
DR   InterPro; IPR001368; TNFR_c6.
DR   SMART; SM00208; TNFR; 1.
DR   PROSITE; PS00652; TNFR_NGFR_1; FALSE_NEG.
DR   PROSITE; PS50050; TNFR_NGFR_2; FALSE_NEG.
KW   Direct protein sequencing; Glycoprotein; Polymorphism; Receptor;
KW   Signal; Transmembrane.
FT   SIGNAL       1     25
FT   CHAIN       26    430       Tumor necrosis factor receptor
FT                               superfamily member 19L.
FT   TOPO_DOM     26    162       Extracellular (Potential).
FT   TRANSMEM    163    183       Potential.
FT   TOPO_DOM    184    430       Cytoplasmic (Potential).
FT   REPEAT       50     90       TNFR-Cys.
FT   CARBOHYD    149    149       N-linked (GlcNAc. . .) (Potential).
FT   DISULFID     51     65       By similarity.
FT   DISULFID     71     90       By similarity.
FT   VARIANT     332    332       A -> V (in dbSNP:12362779).
FT                               /FTId=VAR_022614.
FT   CONFLICT    122    122       D -> S (in Ref. 4).
FT   CONFLICT    187    187       K -> E (in Ref. 2).
FT   CONFLICT    273    273       H -> R (in Ref. 2).
SQ   SEQUENCE   430 AA;  46092 MW;  4A5AB9AE32D36101 CRC64;

  Query Match             21.5%;  Score 307;  DB 1;  Length 430;
  Best Local Similarity   31.0%;  Pred. No. 6e-15;
  Matches   84;  Conservative   36;  Mismatches   69;  Indels   82;  Gaps   12;

Qy          44 TPSPSNDTGNGHPE-----YIAYALVPVPFFIMGLFGVLICHLLKKKGYRCTTEAE----- 93
               |  | | |  | ||     |   |:|||| :||| |:|:|:|||:||| ||   |
Db         144 TRQPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKGYHCTAHKEVGPGP 203

Qy          94 ------------QDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
                         :|                | ||:| :|  | : : ||  |: :: :
Db         204 GGGGSGINPAYRTED---------------ANEDTIGVLVRLITEKKENAAALEELLKE 247

Qy         141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCHRCR 196
                 |  :  |  |     |||  | :|   |  |:|   |||||| |:  :    | ||
Db         248 ---YHSKQLVQTS---HRPVSKLPPAPPNVP--HICPHRHHLHTVQGLASLSGPCCSRCS 299

Qy         197 HKRWHFI-----------------KPTNKSRE-SRPRRQGEVTVLSVGRFRVTKVEHKS 237
                 |:|  :                 ||   :  ::  ||||:|:|||||||| ::
Db         300 QKKWPEVLLSPEAVAATTPVPSLLPNPTRVPKAGAKAGRQGEITILSVGRFRVARI---- 355

Qy         238 NQKERRSLMSVSGAETVN------GEVPATP 262
                 |:|:   ||  :|:       |::| :|
Db         356 --PEQRTSSMVSEVKTITEAGPSWGDLPDSP 384


RESULT 12
TR19L_MOUSE
ID   TR19L_MOUSE    STANDARD;      PRT;   436 AA.
AC   Q8BX43; Q8BTV0;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Tumor necrosis factor receptor superfamily member 19L precursor.
GN   Name=Tnfrsf19l;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C57BL/6J, and NOD; TISSUE=Cerebellum, and Thymus;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
CC   -!- FUNCTION: Mediates activation of NF-kappa-B. May play a role in T-
CC       cell activation (By similarity).
CC   -!- SUBUNIT: Associates with TRAF1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC   -!- SIMILARITY: Contains 1 TNFR-Cys repeat.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AK049073; BAC33534.1; -; mRNA.
DR   EMBL; AK088621; BAC40459.1; -; mRNA.
```

```
DR   HSSP; P19438; 1EXT.
DR   Ensembl; ENSMUSG00000008318; Mus musculus.
DR   MGI; MGI:2443373; Tnfrsf19l.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   InterPro; IPR001368; TNFR_c6.
DR   SMART; SM00208; TNFR; 1.
DR   PROSITE; PS00652; TNFR_NGFR_1; FALSE_NEG.
DR   PROSITE; PS50050; TNFR_NGFR_2; FALSE_NEG.
KW   Glycoprotein; Receptor; Signal; Transmembrane.
FT   SIGNAL        1     31       By similarity.
FT   CHAIN        32    436       Tumor necrosis factor receptor
FT                                superfamily member 19L.
FT   TOPO_DOM     32    169       Extracellular (Potential).
FT   TRANSMEM    170    190       Potential.
FT   TOPO_DOM    191    436       Cytoplasmic (Potential).
FT   REPEAT       57     97       TNFR-Cys.
FT   CARBOHYD    156    156       N-linked (GlcNAc. . .) (Potential).
FT   DISULFID     58     72       By similarity.
FT   DISULFID     78     97       By similarity.
FT   CONFLICT     39     39       G -> C (in Ref. 1; BAC40459).
SQ   SEQUENCE   436 AA;  46472 MW;  1F96C5E021945DF9 CRC64;

  Query Match             20.7%;  Score 295.5;  DB 1;  Length 436;
  Best Local Similarity   30.6%;  Pred. No. 4.5e-14;
  Matches   82;  Conservative   36;  Mismatches   89;  Indels   61;  Gaps   12;

Qy         31 NGSSRTLHSRTETTPSPSNDTGNGHPE-----YIAYALVPVFFIMGLFGVLICHLLKKKG 85
              ||  |           | | |  | ||      |   |:|||| :||| |:|:|:|||:||
Db        148 NGEPR----------QPGNGTRAGGPEETAAQYAVIAIVPVFCLMGLLGILVCNLLKRKG 197

Qy         86 YRCTTEAEQDIEE----EKVEKIELNDSVNENSDTVGQIVHYIMKNEANA----DVLKAM 137
              | || : |          :      :  ||  ||:| :|  | : : ||    ::||
Db        198 YHCTAQKEVGPSPGGGGSGINPAYRTEDANE--DTIGVLVRLITEKKENAAALEELLKEY 255

Qy        138 VADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKHVC--GHHLHTVGGV--VERDVCH 193
             : :    ||    | || :          |:|   |||||| |:  :    |
Db        256 HSKQLVQTSHRPVPRLLPASPSI-----------PHICPHHHHLHTVQGLASLSGPCCS 303

Qy        194 RCRHKRWHFI-----------KPT---NKSR----ESRPRRQGEVTVLSVGRFRVTKVE 234
              ||  | |  :           ||      ||      ::| ||||:|:|||||||| ::
Db        304 RCSQK-WPEVLLSPEAAAATTPAPTLLPTASRAPKASAKPGRQGEITILSVGRFRVARIP 362

Qy        235 HKSNQKERRSLMSVSGAETVNGEVPATP 262
               :       : ::: |    |::| :|
Db        363 EQRTSSLLSEVKTITEAGPSEGDLPDSP 390


RESULT 13
Q8CFT0_MOUSE
ID   Q8CFT0_MOUSE PRELIMINARY;      PRT;   294 AA.
AC   Q8CFT0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   RIKEN cDNA 4631403P03.
GN   Name=4631403P03Rik;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Eye;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038500; AAH38500.1; -; mRNA.
DR   Ensembl; ENSMUSG00000044024; Mus musculus.
DR   MGI; MGI:1918044; 4631403P03Rik.
SQ   SEQUENCE   294 AA;  31456 MW;  8A5DEDE20243EAF6 CRC64;

  Query Match             20.3%;  Score 289;  DB 2;  Length 294;
  Best Local Similarity   38.4%;  Pred. No. 9.1e-14;
  Matches   86;  Conservative   28;  Mismatches   82;  Indels   28;  Gaps    9;

Qy         55 HPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENS 114
              |  |: : || |||:||| | :|||:|||||||| |    : :: :::  | :| |||
Db         13 HGLYMLFLLVLVFFLMGLVGFMICHVLKKKGYRCRTSRGSEPDDAQLQPPE-DDDVNE-- 69

Qy        115 DTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKH 174
              ||| :||  |::|||||: || |: |:   : |  |  |:      |          | |  |
Db         70 DTVERIVRCIIQNEANAEALKEMLGDS---EGEGTVQLSS-----VDATSSLQDGAPSHH 121

Qy        175 VCGHHLHTVGGVVERDVCHRCRHKRWHFIKP-TNKSRESRPRRQGEVTVLSVGRFRVTKV 233
                 |||         |  | ||   ::   :|  :||| | || || |||||||| :
Db        122 ------HTVHLGSAAPCIHCSRSKRPPLVRQGRSKEGKSRP-RPGETTVFSVGRFRVTHI 174

Qy        234 -------EHKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGT 270
                    ||:      ||  |  | |   |   |    : |:||
Db        175 EKRYGLHEHRDGSPTDRSWGSGGGQEP--GGSQAAGGGQPRTGT 216


RESULT 14
Q8BRJ3_MOUSE
ID   Q8BRJ3_MOUSE PRELIMINARY;      PRT;   303 AA.
AC   Q8BRJ3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 10 days neonate cortex cDNA, RIKEN full-length enriched
DE   library, clone:A830089I08 product:hypothetical protein, full insert
DE   sequence.
GN   Name=4631403P03Rik;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Cortex;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Cortex;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
```

```
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Cortex;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Cortex;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Cortex;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Cortex;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK044097; BAC31774.1; -; mRNA.
DR   Ensembl; ENSMUSG00000044024; Mus musculus.
DR   MGI; MGI:1918044; 4631403P03Rik.
KW   Hypothetical protein.
SQ   SEQUENCE   303 AA;  32326 MW;  06287A05AA64793B CRC64;

  Query Match            20.3%;  Score 289;  DB 2;  Length 303;
  Best Local Similarity  38.4%;  Pred. No. 9.4e-14;
  Matches   86;  Conservative   28;  Mismatches   82;  Indels   28;  Gaps    9;

Qy         55 HPEYIAYALVPVFPIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENS 114
              |  |: : || |||:||| | :|||:|||||||| |    : :: :::  | :| |||
Db         13 HGLYMLFLLVLVFFLMGLVGFMICHVLKKKGYRCRTSRGSEPDDAQLQPPE-DDDVNE-- 69

Qy        115 DTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPGKH 174
              ||| :||  |::|||||: || |:\|:   : |  |  |:     |        | |  |
Db         70 DTVERIVRCIIQNEANAEALKEMLGDS---EGEGTVQLSS-----VDATSSLQDGAPSHH 121

Qy        175 VCGHHLHTVGGVVERDVCHRCRHKRWHFIKP-TNKSRESRPRRQGEVTVLSVGRFRVTKV 233
                  |||          |  | ||   ::   :|  :||| | || || ||||||||  :
Db        122 ------HTVHLGSAAPCIHCSRSKRPPLVRQGRSKEGKSRP-RPGETTVFSVGRFRVTHI 174

Qy        234 -------EHKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGT 270
                    ||:      ||  |  | |   |    |    : |:||
Db        175 EKRYGLHEHRDGSPTDRSWGSGGGQEP--GGSQAAGGGQPRTGT 216


RESULT 15
Q6DIW4_XENTR
ID   Q6DIW4_XENTR PRELIMINARY;      PRT;   325 AA.
AC   Q6DIW4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC89179 protein.
GN   Name=MGC89179;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Silurana.
OX   NCBI_TaxID=8364;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Whole body;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC075421; AAH75421.1; -; mRNA.
DR   Ensembl; ENSXETG00000002692; Xenopus tropicalis.
SQ   SEQUENCE   325 AA;  36517 MW;  69129BF0CE1B18CB CRC64;

  Query Match            20.3%;  Score 289;  DB 2;  Length 325;
  Best Local Similarity  35.0%;  Pred. No. 1e-13;
  Matches   82;  Conservative   30;  Mismatches   62;  Indels   60;  Gaps   10;

Qy         58 YIAYALVPVFPIMGLFGVLICHLLKKKGYRCTT-------EAEQDIEEEKVEKIELNDSV 110
              |: : || |||: || | ||||:|||||||| |       ||:: : |:     ||:
Db         13 YMLFLLVLVFFVTGLLGFLICHVLKKKGYRCRTFPEELDPEAKEGLNEDLENAEELS-- 69

Qy        111 NENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTP-------STPGSPPVSPG 163
```

```
                 | ||| :||  |::|||||: || |: |    : | |: |       |    ||
Db         70 --NEDTVERIVKCIIQNEANAEALKQMLGDT---EGEVPIVPSLCLHRESQDAGPP---- 120

Qy        164 PLSPGGTPGKHVCGHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPR-RQGEVTV 222
                          || |||     :  |  |  |:   :    :|:| : :   |||||
Db        121 -------------HH-HTVHLGSTQAPCIHCTRKKRSLLHRMGRSKEGKGKSHPGEVTV 165

Qy        223 LSVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEVPAT--------PVKRERS 268
               |||||||| :  |          |: |::  : | ||:        || |: |
Db        166 FSVGRFRVTHIGKKH---------SIQGSQDGSQE-PASSDKNHEEQPVMRKSS 209
```

Search completed: February 23, 2006, 12:07:49
Job time : 234 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 23, 2006, 12:04:08 ; Search time 40 Seconds
(without alignments)
651.869 Million cell updates/sec

Title: US-10-620-562-2
Perfect score: 1426
Sequence: 1 MAPRALPGSAVLAAAVFVGG.....ETVNGEVPATPVKRERSGTE 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	105.5	7.4	1083	2	JC2300	cell surface glyco
2	104.5	7.3	1679	2	T30271	surface protein -
3	98.5	6.9	3140	2	S47508	genome polyprotein
4	97.5	6.8	1541	2	AG2474	heterocyst glycoli
5	97	6.8	514	2	A31643	cell adhesion 80K
6	96.5	6.8	345	2	S55377	urPAB protein prec
7	96.5	6.8	942	2	JC7316	testicular zinc fi
8	95.5	6.7	634	2	T00359	hypothetical prote
9	94	6.6	682	2	A42121	transcription fact
10	94	6.6	1062	2	T46444	hypothetical prote
11	93.5	6.6	349	2	T25236	hypothetical prote
12	93.5	6.6	452	2	S47633	RXR protein - Afri
13	93.5	6.6	3084	1	MMMSA	laminin alpha-1 ch
14	92.5	6.5	3141	1	GNVSPD	genome polyprotein
15	92	6.5	120	2	T29774	hypothetical prote
16	91.5	6.4	513	2	AB2504	hypothetical prote
17	91.5	6.4	707	2	A46691	E-box-binding prot
18	91.5	6.4	843	2	S33442	EF protein - Strep
19	91.5	6.4	1199	2	T13946	probable adaptor-r
20	91.5	6.4	1822	2	S33441	EF protein - Strep
21	91	6.4	514	2	A44100	cell adhesion mole
22	91	6.4	706	2	S19958	basic helix-loop-h
23	90	6.3	467	2	A43781	retinoid-X-recepto
24	89.5	6.3	612	2	T02414	probable protein k
25	89.5	6.3	2061	2	T13751	transcription fact
26	89	6.2	512	2	T16157	hypothetical prote
27	89	6.2	605	2	S48940	hypothetical prote
28	89	6.2	2364	2	A56577	microtubule-associ
29	89	6.2	3125	1	GNVSPP	genome polyprotein
30	88.5	6.2	563	2	T49681	MCX1 related prote
31	88	6.2	355	2	H90429	membrane conserved
32	88	6.2	441	2	I50515	retinoid X recepto
33	88	6.2	909	2	AG3314	excinuclease ABC c
34	88	6.2	3869	2	A48205	All-1 protein +GTE
35	88	6.2	5107	2	T29144	partial CDS - Caen
36	87.5	6.1	940	2	S19702	fibronectin-bindin
37	87.5	6.1	1420	2	B57062	SRB9 protein - yea
38	87	6.1	890	2	E81576	translation initia
39	87	6.1	890	2	E86530	initiation factor-
40	87	6.1	890	2	F72093	translation initia
41	86.5	6.1	232	2	B96803	hypothetical prote
42	86.5	6.1	879	2	S23006	shed acute-phase a
43	86.5	6.1	916	2	I48921	DNA ligase (ATP) (
44	86	6.0	553	2	B55514	dihydrolipoamide S
45	86	6.0	759	2	T16368	hypothetical prote

ALIGNMENTS

RESULT 1
JC2300
cell surface glycoprotein MSG100 - Pneumocystis carinii
C;Species: Pneumocystis carinii
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: JC2300
R;Wada, M.; Nakamura, Y.
DNA Res. 1, 163-168, 1994
A;Title: MSG gene cluster encoding major cell surface glycoproteins of rat Pneumocystis
A;Reference number: JC2299; MUID:96051989; PMID:8535973
A;Accession: JC2300
A;Molecule type: DNA
A;Residues: 1-1083 <WAD>
A;Cross-references: UNIPROT:Q12075; UNIPARC:UPI000006BAE6; GB:D31909; GB:D17441; NID:g55
C;Genetics:
A;Gene: MSG100
C;Superfamily: Pneumocystis carinii major surface glycoprotein MSG100
C;Keywords: glycoprotein

Query Match 7.4%; Score 105.5; DB 2; Length 1083;
Best Local Similarity 25.3%; Pred. No. 2.5;
Matches 37; Conservative 22; Mismatches 50; Indels 37; Gaps 5;

```
Qy        71 GLFGVLICHLLKKKGYRCTT------EAEQDIEEEKVEKIELNDSVNENSDTVGQIVHYI 124
             |  | :   |::|   || |      :|: :: |:| |  ||  :  |:::    ::
Db       734 GTKGTVCKELVEKLKKRCPTLKDGLNKAKDELTEKKKEYDELKKAAEESTEKAKLLLSKS 793

Qy       125 MK------NEANADVLKAMVADNSLYDPESPVTPSTPGSP----PVSPG---------- 163
               |       ::|         | :    | ||  |  ||||    | :||
Db       794 GKVAMPSGQDSNGSGPVPAPAGSGSGSPSSPAVPPPPGSPQNGTPGTPGGETGASGGTPG 853

Qy       164 -PLSPGGTPGKHVCGHHLHTVGGVVE 188
              | :||| ||         | ||:::
Db       854 TPGTPGGAPG---------TPGGMMK 870
```

RESULT 2
T30271
surface protein - fluke (Schistosoma mansoni)
C;Species: Schistosoma mansoni
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
C;Accession: T30271
R;Hall, T.M.; Joseph, G.T.; Strand, M.
Exp. Parasitol. 80, 242-249, 1995
A;Title: Schistosoma mansoni: molecular cloning and sequencing of the 200-kDa chemothera
A;Reference number: Z20800; MUID:95203407; PMID:7534724
A;Accession: T30271
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1679 <HAL>
A;Cross-references: UNIPROT:Q26607; UNIPARC:UPI0000074F18; EMBL:M99494; NID:g501208; PID

```
  Query Match             7.3%;  Score 104.5;  DB 2;  Length 1679;
  Best Local Similarity   22.8%;  Pred. No. 5.2;
  Matches   51;  Conservative   31;  Mismatches   73;  Indels   69;  Gaps   10;

Qy          79 HLLKKKGYRCTTE-----------------AEQDIEEEKVEKIELNDS--VNENSDTV-- 117
               |||  : |   |                   :  : |:| |  ::| |     :|:| :
Db         634 HLLDIRAYHQLDENNVDSQLKILTLQRSGCSPSESEDEVVASVQLKDGQCSTDNNDIITC 693

Qy         118 -----GQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSPGPLSPGGTPG 172
                  |||: : : | : :|| |  :  : : |       ||      |: | |  | |
Db         694 TRTIHGQIIQFKLNNPSTSDVYKLYMKSDGVEDN----VESTSSIDLVTSGSL--GETVK 747

Qy         173 KHVCG----------HH------------LHTVGGVVERDVCHRCRHKRWHFIKP---TN 207
               : : |          ||            :|    |:  ::  |  :    ||:|   |
Db         748 EDIKGAGLSLTVEGIHHNHETQETELDVAVHIASKVISDNIACRPTYLLLEFIEPNIETL 807

Qy         208 KSRESR---------PRRQGEVTV---LSVGRFRVTKVEHKSNQ 239
               |||  |         |  | |: :   ||:|    |: |  :||
Db         808 KSRVSSKQTMFRIKLPSNQKEINLKMQLSIGSVDPTQSEATTNQ 851


RESULT 3
S47508
genome polyprotein - plum pox virus (isolate PVV-SC)
N;Contains: coat protein; protein 6K1; protein 6K2; protein CI; protein HCpro; protein N
C;Species: plum pox virus, PPV
A;Variety: isolate PVV-SC
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S47508
R;Maiss, E.; Deborre, G.; Jelkmann, W.; Casper, R.
submitted to the EMBL Data Library, August 1994
A;Description: Complete nucleotide sequence of a plum pox potyvirus isolate (PVV-SC) der
A;Reference number: S47508
A;Accession: S47508
A;Molecule type: genomic RNA
A;Residues: 1-3140 <MAI>
A;Cross-references: UNIPROT:Q84925; UNIPARC:UPI00000F3F01; EMBL:X81083; NID:g531731; PID
A;Experimental source: isolate PVV-SC
C;Superfamily: tobacco etch virus genome polyprotein
C;Keywords: ATP; coat protein; genome-linked protein; nucleotide binding; P-loop; phosph
F;1-308/Product: protein P1 #status predicted <P1P>
F;309-766/Product: protein HCpro #status predicted <HCP>
F;767-1116/Product: protein P3 #status predicted <P3P>
F;1117-1168/Product: protein 6K1 #status predicted <6K1>
F;1169-1803/Product: protein CI #status predicted <CIP>
F;1253-1260/Region: nucleotide-binding motif A (P-loop)
F;1338-1343/Region: nucleotide-binding motif B
F;1342-1345/Region: DEXH motif
F;1804-1856/Product: protein 6K2 #status predicted <6K2>
F;1857-2292/Product: protein NIa #status predicted <NIA>
F;1857-2049/Product: VPg protein #status predicted <VPG>
F;2293-2810/Product: protein NIb #status predicted <NIB>
F;2811-3140/Product: coat protein #status predicted <COA>
F;1919/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

  Query Match             6.9%;  Score 98.5;  DB 2;  Length 3140;
  Best Local Similarity   24.7%;  Pred. No. 35;
  Matches   45;  Conservative   27;  Mismatches   85;  Indels   25;  Gaps    6;

Qy          83 KKGYRCTTEAEQDIEEE-KVEKIELNDSVNENSDTVGQIVHYIMKNEANA---DVLKAMV 138
               || |  |  :| :||   |    ::||   |::  | |:     : | ||   :|: |
Db        2774 KKLYTDTEASEIEIERYLKAFYDDINDD-GESNVVVHQVAEREDEEEVNAGKSNVVTAPA 2832

Qy         139 ADNSLYDP-----ESPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHLHTVGGVVERDVCH 193
               | : :  |       :| | ::  :|  :|    |    | | | | | |     |
Db        2833 ATSPILQPPLVIQPAPRTTASMLNPIFTPATTQPATKPVSQVSGPQLQTFGTYGNEDA-- 2890

Qy         194 RCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLMSVSGAET 253
                     |:| :      |  :|   | | | | :::  ::   : ||  | |
Db        2891 ----------SPSNSNALVNTSRDRDVDAGSTGTFTVPRLKAMTS---KLSLPKVKGKAI 2937

Qy         254 VN 255
               :|
Db        2938 MN 2939


RESULT 4
AG2474
heterocyst glycolipid synthase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AG2474
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG2474
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1541 <KUR>
A;Cross-references: UNIPROT:Q8YLE7; UNIPARC:UPI00000CEE52; GB:BA000019; PIDN:BAB77050.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: hglE

  Query Match             6.8%;  Score 97.5;  DB 2;  Length 1541;
  Best Local Similarity   21.3%;  Pred. No. 17;
  Matches   45;  Conservative   30;  Mismatches   73;  Indels   63;  Gaps    8;

Qy           1 MAPRA--LPGSAVLAAAVFVGG------AVSSPLVAPDNGSSRTLHSRTETTPSPSNDTG 52
               :|| |  :    |:||  |          : :|: ||       ::   | | :|
Db        1248 LAPTAWQIQSEPVVAATASVVSTIEPEPVIEAPVAAP------VVNQVVEFTSAP----- 1296

Qy          53 NGHPEYIAYALVPVFFIMGLFGVLICHLLKKKGY-----RCTTEAEQDIEEEKVEKIELN 107
                     |    |   :  |   |:  :  | ||         : | |:  : ::::|:
Db        1297 -------APVAAPAIDVADLDKNLLAIISDKTGYPVEMLELEMDMEADLGIDSIKRVEIL 1349

Qy         108 DSVNE-------------NSDTVGQIVHYIMKNEA------------NADVLKAMVAD 140
                 : |                 |:|||| |:  | |              |  : | |
Db        1350 GGLQEVYPDLPKPNLEELAEKRTIGQIVDYLKANTAISQSVGIAIQQAQQAAEIPAPVVT 1409

Qy         141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGTP 171
                 |:  ||  | |:     |::| |::  ||
Db        1410 ASIAAPEPVVAPT-----PIAPEPVAIPATP 1435


RESULT 5
A31643
cell adhesion 80K glycoprotein precursor - slime mold (Dictyostelium discoideum)
N;Alternate names: cell adhesion molecule csA; cell surface glycoprotein, 80K
C;Species: Dictyostelium discoideum
A;Variety: strain AX-2
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: S22066; A31643; A26310; A23951; A33463; S01473
R;Desbarats, L.; Lam, T.Y.; Wong, L.M.; Siu, C.H.
submitted to the EMBL Data Library, June 1992
A;Description: Identification of a unique camp-responsive element in the gene encodidg th
A;Reference number: S22066
A;Accession: S22066
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-514 <DES>
A;Cross-references: UNIPROT:P08796; UNIPARC:UPI00001284F3; EMBL:X66483; NID:g7289; PIDN:(
R;Kamboj, R.K.; Wong, L.M.; Lam, T.Y.; Siu, C.H.
J. Cell Biol. 107, 1835-1843, 1988
A;Title: Mapping of a cell-binding domain in the cell adhesion molecule gp80 of Dictyoste
A;Reference number: A31643; MUID:89034443; PMID:3182938
A;Accession: A31643
A;Molecule type: mRNA
A;Residues: 1-514 <KAM>
A;Cross-references: UNIPARC:UPI00001284F3; GB:M36545; NID:g167803; PIDN:AAA33212.1; PID:g
R;Noegel, A.; Gerisch, G.; Stadler, J.; Westphal, M.
EMBO J. 5, 1473-1476, 1986
A;Title: Complete sequence and transcript regulation of a cell adhesion protein from aggr
```

```
A;Reference number: A26310
A;Accession: A26310
A;Molecule type: mRNA
A;Residues: 1-215,'G',217-514 <NOE>
A;Cross-references: UNIPARC:UPI000016B99B; EMBL:X04004; NID:g7246; PIDN:CAA27634.1; PID:
R;Wong, L.M.; Siu, C.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 4248-4252, 1986
A;Title: Cloning of cDNA for the contact site A glycoprotein of Dictyostelium discoideum
A;Reference number: A23951
A;Accession: A23951
A;Molecule type: mRNA
A;Residues: 20-49 <WON>
A;Cross-references: UNIPARC:UPI0000000031; GB:M13546; NID:g167699; PIDN:AAA33181.1; PID:
R;Kamboj, R.K.; Gariepy, J.; Siu, C.H.
Cell 59, 615-625, 1989
A;Title: Identification of an octapeptide involved in homophilic interaction of the cell
A;Reference number: A33463; MUID:90058644; PMID:2582489
A;Accession: A33463
A;Molecule type: protein
A;Residues: 132-139 <KA2>
A;Cross-references: UNIPARC:UPI000017B17A
R;Stadler, J.; Bordier, C.; Lottspeich, F.; Henschen, A.; Gerisch, G.
Hoppe-Seyler's Z. Physiol. Chem. 363, 771-776, 1982
A;Title: Improved purification and N-terminal amino acid sequence determination of the c
A;Reference number: S01473; MUID:83005404; PMID:7118072
A;Accession: S01473
A;Molecule type: protein
A;Residues: 20-37,'X',39-41,'X',43-46 <STA>
A;Cross-references: UNIPARC:UPI000017B17B
R;Stadler, J.; Keenan, T.W.; Bauer, G.; Gerisch, G.
EMBO J. 8, 371-377, 1989
A;Title: The contact site A glycoprotein of Dictyostelium discoideum carries a phospholi
A;Reference number: A56857; MUID:89251561; PMID:2721485
A;Contents: annotation; detection of glycosphingolipidinositol anchor
A;Note: no phosphoserine was detected but phosphoinositol and ceramide were
A;Note: the N-linked carbohydrates are sulfated type 1 oligosaccharides
C;Comment: This protein is involved in the formation of intercellular contacts upon aggr
C;Genetics:
A;Gene: csA
C;Keywords: blocked carboxyl end; cell adhesion; glycoprotein; lipoprotein; membrane pro
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-514/Product: cell adhesion 80K glycoprotein #status predicted <MAT>
F;132-139/Region: cell adhesion #status predicted
F;460-479/Region: 10-residue repeats
F;493-514/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;128,137,207,294,399/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;492/Modified site: GSI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form

  Query Match             6.8%;  Score 97;  DB 2;  Length 514;
  Best Local Similarity   19.6%;  Pred. No. 4.8;
  Matches   57;  Conservative   50;  Mismatches  116;  Indels   68;  Gaps   13;

Qy         16 VFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPSN------------DTGNGHPEYIAYAL 63
              | :||    |::  :  | :   |  :  |  ||              |||  :|    |
Db         50 VTIGGQTCDPVIVANTASLQCQFS-AQLAPGNSNFDVIVKVGGVPSTGGNGLFKYTPPTL 108

Qy         64 VPVFFIMGLFG-VLICHLLKKKGYRCTTEAEQDIEEEKVEKIELNDSVNE-----NSDTV 117
                :|   |  | :|:       ||              |: :|||:|        :|:|
Db        109 STIFPNNGRIGMILVDGPSNISGY----------------KLNVNDSINSAMLSVTADSV 152

Qy        118 GQIVHYIMKNEANADVL----------KAMVADNSLYDPESPVTPSTPGSPPVSPGPLSP 167
                 ::::: |     :|           :|   |:: |   :|  || :  ::|  ::
Db        153 SPTIYFLVPNTIAGGLLNLELIQPFGFSTIVTSKSVFSP--TITSITPLAFDLTPTNVTV 210

Qy        168 GG----TPGKHVCGHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVL 223
               |    |      | |::|  |:  :|     |     | | |   ||       :|
Db        211 TGKYFVTTASVTMGSHIYT--GLTVQDDGTNC-----HVIFTTRSVYES----SNTITAK 259

Qy        224 SVGRFRVTKVEHKSNQKE---RRSLMSVSGAETVNGEVPATPVKRERSGTE 271
               :    :  :::: ||:     :  :::  : ||  |    :    :||:
Db        260 ASTGVDMIYLDNQGNQQPITFTYNPPTITSTKQVNDSV---EISTTNTGTD 307
```

```
RESULT 6
S55377
urPAB protein precursor - Peptostreptococcus magnus
C;Species: Peptostreptococcus magnus
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S55377
R;de Chateau, M.; Bjoerck, L.
submitted to the EMBL Data Library, April 1995
A;Description: Protein urPAB.
A;Reference number: S55377
A;Accession: S55377
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-345 <DEC>
A;Cross-references: UNIPROT:Q51932; UNIPARC:UPI00000AE87E; EMBL:Z48975; NID:g854370; PID

  Query Match             6.8%;  Score 96.5;  DB 2;  Length 345;
  Best Local Similarity   20.3%;  Pred. No. 3.2;
  Matches   47;  Conservative   17;  Mismatches   64;  Indels  103;  Gaps    7;

Qy         97 EEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDP---------- 146
              |||| :|    |  :   |        |  | ||     ||:  :| | :|
Db        121 EEEKAKKEAEKDVSDSWYDQSKHPFAYRTKEEAEQAAKKAL--ENELLNPGHINNSYKVS 178

Qy        147 ----------------ESPVTPSTPGSPPVSPGPLSP---------------------- 167
                             |:| ||  |  | |:| | :|
Db        179 QKYDGNWEYVLSPNLAETPETPKKPSKPDVTPNPSTPEVEEKDPVLPSYPRLDNHYFMNG 238

Qy        168 -------GGTPGKHVCGHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEV 220
                  | ||||                                ||    :| :| ::
Db        239 DSSTSDEGKTPGKEE---------------------------KP---GKEEKPGKE--- 264

Qy        221 TVLSVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGTE 271
                     | |  ::|:       | |      || | | |:  |:
Db        265 ----------EKPEQKPGKEEKPGKEEKPGKEEKPAN-PAKPAKEEKEKTD 304
```

```
RESULT 7
JC7316
testicular zinc finger protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
C;Accession: JC7316
R;Inoue, A.; Ishiji, A.; Kasagi, S.; Ishizuka, M.; Hirose, S.; Baba, T.; Hagiwara, H.
Biochem. Biophys. Res. Commun. 273, 398-403, 2000
A;Title: The transcript for a novel protein with a zinc finger motif is expressed at spe
A;Reference number: JC7316
A;Contents: Testis
A;Accession: JC7316
A;Molecule type: mRNA
A;Residues: 1-942 <INO>
A;Cross-references: UNIPROT:Q9JJ01; UNIPARC:UPI00000285D3
C;Comment: This protein, with a zinc finger motif, functions as a transcriptional regula
in the initiation of the reduction divisions of meiosis and/or in the maintenance of mei
C;Genetics:
A;Gene: tzf
C;Keywords: meiosis; spermatogenesis; testis; transcription regulation; zinc finger

  Query Match             6.8%;  Score 96.5;  DB 2;  Length 942;
  Best Local Similarity   21.4%;  Pred. No. 11;
  Matches   75;  Conservative   39;  Mismatches  134;  Indels  103;  Gaps   13;

Qy          5 ALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETTPSPS---------NDTGNGH 55
              ::|| | :      | ::  |   |        |  | |    ||          :|  |
Db        172 SIPGLAGVLTTSEPGYSLQRPEEVPMMPKKSILKKRIEADMKPSLQLESFSSGASSGEDH 231

Qy         56 PEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIE---LNDSVNE 112
              | |  :: :|:   :  |       :: ||     |  | |::| :    :   | :  :
Db        232 PLYSEHSPLPLSGAIAAP----TSEIENKG----TTVEADLKEPQSNLYQWGPLREIPKD 283

Qy        113 NSDTVGQIVHYIMKNEANADVLK----------------------AMVAD--------- 140
              ||:     : :  : :  |: |:                        :|:||
Db        284 NSEKFDSFLGPKERLDLKAEGLEQQTDNLLPHERASQDGSGFSRILSMLADPTITQEKRR 343
```

```
Qy        141 --------------------------NSLYDPESPVT----PSTPGSPPVSPGPLSPGG 169
                                         || ||||  |     | | :| |    |
Db        344 RSFPDIEDEEKFLYGDEEEDIKSESPLKSLEDPESAGTRQKANSLPSTPAVKLESLEE-S 402

Qy        170 TPGKHVCGHHLHTVG---GVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVG 226
               |      : | |:|   || |         :| |  ||:  || |  ||      ||
Db        403 NPEYAKIHNLLKTIGLDIGVAEIGKLAARTQERLHGKKPS--SRPSADRRLSADRHLSGD 460

Qy        227 RF----RVTKVEH------KSNQKERRSLMSVSGAETVNGEV----PATPV 263
               |     | : |||      :|:   |        : | : ||    ||:||
Db        461 RHFSADRCSSVEHSFTADWRSSDPHRPESRETHHSNTQSPEVSHPHPASPV 511


RESULT 8
T00359
hypothetical protein KIAA0680 - human
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00359
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00359
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-634 <ISH>
A;Cross-references: UNIPROT:O75167; UNIPARC:UPI000006E0AA; EMBL:AB014580; NID:g3327173;
A;Experimental source: brain; clone HK02746
C;Genetics:
A;Note: KIAA0680

  Query Match             6.7%;  Score 95.5;  DB 2;  Length 634;
  Best Local Similarity   19.5%;  Pred. No. 8.2;
  Matches   52;  Conservative   24;  Mismatches   79;  Indels  111;  Gaps   10;

Qy         50 DTGNGHPEYIAYALVPVPFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEKIEL-ND 108
              :  |||      ::|:                          |:   || | | |  |
Db         96 ENSNGH-------MIPI------------------------GEESTREENVVKSEEGNG 123

Qy        109 SVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPVSP------ 162
              ||:| :  :          |  |:  |    :|:    |:|  |: | | |  |
Db        124 SVSEKTPPL----------EEQAEDKK----ENTENHSETPAAPALPPSAPPKPRSKPKP 169

Qy        163 --GPLSPGG--------------------TPGKHV-----------------CGHHLHT 182
                |: | |                    |||                     |   |
Db        170 KKSPVPPKGATAGASHKGDEVPPIKKNTKAPGKQAPVPPPKPASRNTTREAAGSSHSKKT 229

Qy        183 VGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQKER 242
               |              |:  |  |||:   |      |   || : |::::
Db        230 TGSKAS-------------ASPSTSSTSSRPKASKETVSSKAGTVGTTKGKRKTDKQPI 275

Qy        243 RSLMS----VSGAETVNGEVPATPVK 264
               | :|     ||   : ||   | |:
Db        276 TSHLSSDTTTSGTSDLKGEPAETRVE 301


RESULT 9
A42121
transcription factor HTF4 - human
N;Alternate names: helix-loop-helix protein (HEB)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A42121; B56611; S34423; S34424
R;Hu, J.S.; Olson, E.N.; Kingston, R.E.
Mol. Cell. Biol. 12, 1031-1042, 1992
A;Title: HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the D
A;Reference number: A42121; MUID:92186835; PMID:1312219
A;Accession: A42121
A;Molecule type: mRNA
A;Residues: 1-682 <HU1>
A;Cross-references: UNIPROT:Q99081; UNIPARC:UPI000003B45E; GB:M80627; NID:g183929; PIDN:
A;Note: sequence extracted from NCBI backbone (NCBIN:88251, NCBIP:88254)
R;Zhang, Y.; Bina, M.
DNA Seq. 2, 397-403, 1992
A;Title: The nucleotide sequence of the human transcription factor HTF4a cDNA.
A;Reference number: A56611; MUID:93076001; PMID:1446075
A;Accession: B56611
A;Molecule type: mRNA
A;Residues: 1-682 <ZHA>
A;Cross-references: UNIPARC:UPI000003B45E; GB:M83233; NID:g2243159; PIDN:AAB62389.1; PID
A;Note: sequence extracted from NCBI backbone (NCBIN:118868, NCBIP:118873)
A;Note: it is uncertain whether Met-1, Met-8, or Met-26 is the initiator
R;Zhang, Y.; Babin, J.; Feldhaus, A.L.; Singh, H.; Sharp, P.A.; Bina, M.
Nucleic Acids Res. 19, 4555, 1991
A;Title: HTF4: a new human helix-loop-helix protein.
A;Reference number: S34423; MUID:91360364; PMID:1886779
A;Accession: S34423
A;Molecule type: mRNA
A;Residues: 570-655 <ZH2>
A;Cross-references: UNIPARC:UPI0000176417; EMBL:M65209
C;Genetics:
A;Gene: GDB:TCF12
A;Cross-references: GDB:355444; OMIM:600480
A;Map position: 15q21-15q21
C;Superfamily: human transcription factor 3
C;Keywords: DNA binding; transcription factor

  Query Match             6.6%;  Score 94;  DB 2;  Length 682;
  Best Local Similarity   20.6%;  Pred. No. 12;
  Matches   73;  Conservative   38;  Mismatches  109;  Indels  134;  Gaps   16;

Qy          3 PRALPGSAVLAAAVFV--GGAVSSPLVAPDNGSSRT----LHSRTETTPSPSNDTGNGH- 55
              |   | ::: |:  |:  |   || | :  || |:     :  : :  | |:  || |
Db        212 PSPKPPTSMFASTFFMQDGTHNSSDLWSSSNGMSQPGFGGILGTSTSHMSQSSSYGNLHS 271

Qy         56 ------------PEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVEK 103
                         |  |  :| |    | |           :| :| :   :
Db        272 HDRLSYPPHSVSPTDINTSLPP----MSSF-----------HRGSTSSSPYVAASHTPP 315

Qy        104 IELNDSV----------NENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPE--SPVT 151
              |  :||:          ::  | :|                 ||:    |:| |:  |
Db        316 INGSDSILGTRGNAAGSSQTGDALG---------------KALA---SIYSPDHTSSSF 356

Qy        152 PSTPGSPPVSPGPLS-------PGG-TPGKHVCGHHLHTVGGVVE------RDVCHRCRH 197
              || | :|  || ||:       |||  |      : ||::  :|       | |  |:
Db        357 PSNPSTPVGSPSPLTGTSQWPRPGGQAPSSPSYENSLHSLQSRMEDRLDRLDDAIHVLRN 416

Qy        198 KRWHFIKPTNK----------------------------------------------SRESR 213
                | : |:                                              ||
Db        417 ---HAVGPSTSLPAGHSDIHSLLGPSHNAPIGSLNSNYGGSSLVASSRSASMVGTHREDS 473

Qy        214 PRRQGEVTVL-SVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEVPATPVKRE 266
                  :|| |      | : ||: :  |  | | ||       |  | :| |
Db        474 VSLNGNHSVLSSTVTTSSTDLNHKTQENYRGGLQSQSGT------VVTTEIKTE 521


RESULT 10
T46444
hypothetical protein DKFZp434N1427.1 - human
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46444
R;Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23032
A;Accession: T46444
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1062 <AAA>
A;Cross-references: UNIPROT:Q9NSW2; UNIPARC:UPI000006FE55; EMBL:AL137701
A;Experimental source: adult testis; clone DKFZp434N1427
C;Genetics:
A;Note: DKFZp434N1427.1
```

```
  Query Match             6.6%;  Score 94;  DB 2;  Length 1062;
  Best Local Similarity   26.0%;  Pred. No. 21;
  Matches   45;  Conservative   34;  Mismatches   68;  Indels   26;  Gaps   12;

Qy         91 EAEQDIEEEKVEKIELND--SVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLY-DPE 147
              | |:: |||: |  |  :   : |   ||           ::: ::  :   ::|| ||
Db        208 EEEEEEEEEEEEGPEAREMAMMQEGEHTV----------TSHSSIIHRLPGSDNLYDDPY 257

Qy        148 SP-VTPSTPGSPPVSPGPLSPGGTPGKHV-CGH--HLHTV--GGVVERDVCHRCRHKRWH 201
               | :||| |  || :| | |   :  :   | :  |  |:    :| | :     |::
Db        258 QPEITPS-PLQPPAAPAPTSTTSSARRRAYCRNRDHFATIRTASLVSRQI---QEHEQDS 313

Qy        202 FIKPTNKSRESRPRRQGEVTVLSV-GRFRVTKVEHKSN-QKERRSLMSVSGAE 252
               ::    |   | ||| :  :|::  | |  : || :  |:|  :  :  |||
Db        314 ALR-EQLSGYKRMRRQHQKQLLALESRLRGEREEHSARLQRELEAQRAGFGAE 365


RESULT 11
T25236
hypothetical protein T24D1.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25236
R;Cummings, P.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20001
A;Accession: T25236
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-349 <WIL>
A;Cross-references: UNIPROT:O02333; UNIPARC:UPI000007ED28; EMBL:Z81131; PIDN:CAB03425.1;
A;Experimental source: clone T24D1
C;Genetics:
A;Gene: CESP:T24D1.3
A;Map position: 1
A;Introns: 29/2; 298/2

  Query Match             6.6%;  Score 93.5;  DB 2;  Length 349;
  Best Local Similarity   19.8%;  Pred. No. 5.7;
  Matches   41;  Conservative   32;  Mismatches   67;  Indels   67;  Gaps    9;

Qy         92 AEQDIEEEKVEKIELNDSVNENSDTVGQIVH----YIMKNEANADVLKAMVADNSLYDPE 147
              :| |:|:::   :|: :    |     |::|     : |:||                |
Db         21 SEVDLEQQR-RALEMFEQ-QMNDQAFAQLIHDHDVQHINNDAN-------------QVPA 65

Qy        148 SPVTPSTPGSPPVSPGPLSPGGTPGKHVCGHHLHTVGGVVER------------DVCHRC 195
              :| ||  |  |||     ||  :|             |: |              |: |
Db         66 APNTPPVPVPPPVQQPQASPNASP------------VSVLRRRRYANPEQERFGDISRRI 113

Qy        196 RHKRWHFI---------------KPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQK 240
               |  : |               :| : |:|       |:|: | |   :        ::
Db        114 -HNPANQIPAAPNTQPEVPPTVQQPESSSQERASPTSPEITIPSTGLLGL--------RR 164

Qy        241 ERRSLMSVSGAETVNGEVPATPVKRER 267
                 : :|:  : |    :| ||::| |
Db        165 RSSARISIRMSATAQTNLPITPIRRRR 191


RESULT 12
S47633
RXR protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S47633
R;Marklew, S.; Smith, D.P.; Mason, C.S.; Old, R.W.
Biochim. Biophys. Acta 1218, 267-272, 1994
A;Title: Isolation of a novel RXR from Xenopus that most closely resembles mammalian RXR
A;Reference number: S47633; MUID:94325332; PMID:8049252
A;Accession: S47633
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-452 <MAR>
A;Cross-references: UNIPROT:Q91613; UNIPARC:UPI00000FD9F4; GB:S73269; NID:g685084; PIDN:

A;Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 118-His
C;Superfamily: retinoic acid receptor alpha; erbA transforming protein homology
C;Keywords: zinc finger
F;126-368/Domain: erbA transforming protein homology <ERBA>

  Query Match             6.6%;  Score 93.5;  DB 2;  Length 452;
  Best Local Similarity   23.0%;  Pred. No. 7.8;
  Matches   60;  Conservative   37;  Mismatches   97;  Indels   67;  Gaps   15;

Qy         28 APDNGS-SRTL-HSRTETTPSPSNDTGNGHPEYIAYALV-----PVFFIMGLFGVLICHL 80
              :||  | |  | || ::| | ||      ||  :  |:      |:  |   | |: |
Db          9 SPDTSSLSPPLGHSFSDTPPPPSAPL---HPSMVGSAMTSSVNSPLGSIGSPFPVINCS- 64

Qy         81 LKKKGYRCTTEAEQDIEEEKVEKIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVAD 140
              :  |  |      |    |   ::| :|| :       :|::  :|
Db         65 VGSPGIPGT----PSIGYGPVSSPQINSTVNLSG------LHHVGSSE------------ 102

Qy        141 NSLYDPESPVTPSTPGSPPVSPGPLSPGGT-PGKHVC--------GHH--LHTVGG---- 185
                 | :|:   :  | |         |||  || :|         | |  :|:  |
Db        103 ----DVKPPLGMRSMQSHP-------NGGTVSGKRLCAICGDRSSGKHYGVHSCEGCKGF 151

Qy        186 ---VVERDVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQKER 242
                : :|: : ||  :   :    ::|    | |      |: |  |    | :   :||
Db        152 FKRTIRKDLTYTCRDSKDCIVDKRQRNRCQYCRYQ---KCLATGMKREAVQEERQRGRER 208

Qy        243 RSLMSVSGAETVNGEVPATPV 263
                  :|||  :| |:|   :
Db        209 DGEAELSGA--INEEMPVEKI 227


RESULT 13
MMMSA
laminin alpha-1 chain precursor - mouse
N;Alternate names: laminin chain A1
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C;Accession: A31771; A30449; S00624; A30450; S08895; S02678; S01790; A30451; S14670
R;Sasaki, M.; Kleinman, H.K.; Huber, H.; Deutzmann, R.; Yamada, Y.
J. Biol. Chem. 263, 16536-16544, 1988
A;Title: Laminin, a multidomain protein. The A chain has a unique globular domain and ho
A;Reference number: A31771; MUID:89034134; PMID:3182802
A;Accession: A31771
A;Molecule type: mRNA
A;Residues: 1-3084 <SAS>
A;Cross-references: UNIPROT:P19137; UNIPARC:UPI00000278CA; EMBL:J04064; NID:g309419; PID
A;Accession: A30449
A;Molecule type: protein
A;Residues: 183-195;570-571,'A',573-586;596-612,'X',614-617,'EMK';630-646;1217-1222,'YPF
2486;2624-2639;2818-2843;3009-3033,'V',3035 <SA2>
A;Cross-references: UNIPARC:UPI0000173C92; UNIPARC:UPI0000173C93; UNIPARC:UPI0000173C94;
C99; UNIPARC:UPI0000173C9A; UNIPARC:UPI0000173C9B; UNIPARC:UPI0000173C9C; UNIPARC:UPI000
I0000173CA2
R;Hartl, L.; Oberbaeumer, I.; Deutzmann, R.
Eur. J. Biochem. 173, 629-635, 1988
A;Title: The N terminus of laminin A chain is homologous to the B chains.
A;Reference number: S00624; MUID:88225080; PMID:3267223
A;Accession: S00624
A;Molecule type: mRNA
A;Residues: 1-208,'T',210-334 <HAR>
A;Cross-references: UNIPARC:UPI000016CE8F; EMBL:X07737; NID:g52857; PIDN:CAA30561.1; PID
A;Accession: A30450
A;Molecule type: protein
A;Residues: 311-335,'N',337-339;630-642,'D',644;692-734;737-748,'X',750-760,'G',762-763;
3-1389;1449-1459 <HA2>
A;Cross-references: UNIPARC:UPI0000173CA3; UNIPARC:UPI0000173CA4; UNIPARC:UPI0000173CA5;
CAA; UNIPARC:UPI0000173CAB; UNIPARC:UPI0000173CAC
A;Note: the sequence from Fig. 7 is inconsistent with that from Fig. 5 in having 209-Ile
R;Mann, K.; Deutzmann, R.; Timpl, R.
Eur. J. Biochem. 178, 71-80, 1988
A;Title: Characterization of proteolytic fragments of the laminin-nidogen complex and th
A;Reference number: S08895; MUID:89078415; PMID:2462498
A;Accession: S08895
A;Molecule type: protein
```

```
A;Residues: 153-169 <MAN>
A;Cross-references: UNIPARC:UPI0000173CAD
R;Fujiwara, S.; Shinkai, H.; Deutzmann, R.; Paulsson, M.; Timpl, R.
Biochem. J. 252, 453-461, 1988
A;Title: Structure and distribution of N-linked oligosaccharide chains on various domain
A;Reference number: S02678; MUID:88326259; PMID:2458101
A;Accession: S02678
A;Molecule type: protein
A;Residues: 630-642,'D',644;2690-2704 <FUJ>
A;Cross-references: UNIPARC:UPI0000173CA4; UNIPARC:UPI0000173CAE
R;Deutzmann, R.; Huber, J.; Schmetz, K.A.; Oberbaeumer, I.; Hartl, L.
Eur. J. Biochem. 177, 35-45, 1988
A;Title: Structural study of long arm fragments of laminin. Evidence for repetitive C-te
A;Reference number: S01790; MUID:89030693; PMID:3181157
A;Accession: S01790
A;Molecule type: mRNA
A;Residues: 2538-3084 <DEU>
A;Cross-references: UNIPARC:UPI000016CE8E; EMBL:X13459; NID:g55499; PIDN:CAA31807.1; PID
A;Accession: A30451
A;Molecule type: protein
A;Residues: 1911-1929;1997-2006;2033-2045,'X',2047-2054,'X',2056-2066,'X',2068-2105;2120
470;2487-2498;2502-2525;2538-2557;2561-2591,'X',2593-2594;2600-2610;2616-2645;2648-2655;
93;2998-3005,'A',3007-3033,'V',3035;3068-3083 <DE2>
A;Cross-references: UNIPARC:UPI0000173CB6; UNIPARC:UPI0000173CB7; UNIPARC:UPI0000173CB8;
CBD; UNIPARC:UPI0000173CBE; UNIPARC:UPI0000173CBF; UNIPARC:UPI0000173CC0; UNIPARC:UPI000
I0000173CC6; UNIPARC:UPI0000173CC7; UNIPARC:UPI0000173CC8; UNIPARC:UPI0000173CC9; UNIPAR
A;Note: 2256-Val was also found
R;Olsen, D.; Nagayoshi, T.; Fazio, M.; Peltonen, J.; Jaakkola, S.; Sanborn, D.; Sasaki,
Lab. Invest. 60, 772-782, 1989
A;Title: Human laminin: cloning and sequence analysis of cDNAs encoding A, B1 and B2 cha
A;Reference number: A34961; MUID:89280632; PMID:2733383
A;Accession: S14670
A;Molecule type: protein
A;Residues: 2424-2436;2440-2451;2461-2467;2487-2525;2550-2557;2561-2593;2600-2610;2616-2
-2942,'T',2944-2964;2969-2976;2980-2993;2998-3000,'I',3002-3018,'V',3020-3034;3068-3083
A;Cross-references: UNIPARC:UPI0000173C9E; UNIPARC:UPI0000173CAE; UNIPARC:UPI0000173CBC;
CC7; UNIPARC:UPI0000173CC8; UNIPARC:UPI0000173CCA; UNIPARC:UPI0000173CCB; UNIPARC:UPI000
I0000173CD1; UNIPARC:UPI0000173CD2; UNIPARC:UPI0000173CD3; UNIPARC:UPI0000173CD4; UNIPAR
C;Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin
C;Function:
A;Description: interact with cells and with other basement membrane proteins to promote
C;Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like h
C;Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-3084/Product: laminin alpha-1 chain #status predicted <MAT>
F;25-277/Domain: VI <DOM6>
F;277-331/Domain: laminin-type EGF-like homology <LE01>
F;278-519/Domain: V <DOM5>
F;334-401/Domain: laminin-type EGF-like homology <LE02>
F;404-458/Domain: laminin-type EGF-like homology <LE03>
F;461-507/Domain: laminin-type EGF-like homology <LE04>
F;510-519/Domain: laminin-type EGF-like homology #status atypical <LE05>
F;520-715/Domain: IVb <DO4B>
F;716-1166/Domain: IIIb <DO3B>
F;716-746/Domain: laminin-type EGF-like homology #status atypical <LE06>
F;749-795/Domain: laminin-type EGF-like homology <LE07>
F;798-853/Domain: laminin-type EGF-like homology <LE08>
F;830-834/Region: cell adhesion #status predicted
F;856-906/Domain: laminin-type EGF-like homology <LE09>
F;909-955/Domain: laminin-type EGF-like homology <LE10>
F;958-1002/Domain: laminin-type EGF-like homology <LE11>
F;1005-1048/Domain: laminin-type EGF-like homology <LE12>
F;1051-1094/Domain: laminin-type EGF-like homology <LE13>
F;1097-1116/Domain: laminin-type EGF-like homology #status atypical <LE14>
F;1118-1154/Domain: laminin-type EGF-like homology #status atypical <LE15>
F;1147-1149/Region: cell attachment (R-G-D) motif
F;1157-1166/Domain: laminin-type EGF-like homology #status atypical <LE16>
F;1167-1368/Domain: IVa <DO4A>
F;1369-1561/Domain: IIIa <DO3A>
F;1369-1407/Domain: laminin-type EGF-like homology #status atypical <LE17>
F;1410-1456/Domain: laminin-type EGF-like homology <LE18>
F;1459-1513/Domain: laminin-type EGF-like homology <LE19>
F;1516-1560/Domain: laminin-type EGF-like homology <LE20>
F;1562-2133/Domain: II/I <DOM2>
F;1562-2133/Region: heptad repeats
F;2134-3084/Domain: G <DOMG>
F;2150-2308/Domain: laminin G repeat homology <LG1>
F;2337-2492/Domain: laminin G repeat homology <LG2>
F;2518-2683/Domain: laminin G repeat homology <LG3>
F;2748-2897/Domain: laminin G repeat homology <LG4>
F;2925-3082/Domain: laminin G repeat homology <LG5>
F;25/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;45,79,370,374,531,562,672,808,914,959,969,1052,1344,1414,1586,1603,1659,1686,1706,1718
e (Asn) (covalent) #status predicted
F;304-312/Disulfide bonds: #status experimental
F;770,857,1999,2055,2067,2835/Binding site: carbohydrate (Asn) (covalent) #status experim
F;845,2102/Binding site: carbohydrate (Asn) (covalent) #status absent

  Query Match             6.6%;  Score 93.5;  DB 1;  Length 3084;
  Best Local Similarity   22.2%;  Pred. No. 85;
  Matches   69;  Conservative   33;  Mismatches   94;  Indels  115;  Gaps   16;

Qy         60 AYALVPVFFIMGLFGVLICHL-------LKKKGYRCTTEAEQDIEEEKVEKIE----LND 108
              | | || | || | | :  |:         |:|      | :  | :|  :  :     :
Db       2545 AQAHVPFFSIMLLEGRIEVHVNSGDGTSLRKALLHAPTGSYSDGQEHSISLVRNRRVITI 2604

Qy        109 SVNENSD--------TVGQIVH----YI-------------MKNEANADVLKAMVADNSL 143
               |:|||          | |: :     ||             |:   :   :| :| |  |
Db       2605 QVDENSPVEMKLGPLTEGKTIDISNLYIGGLPEDKATPMLKMRTSFHG-CIKNVVLDAQL 2663

Qy        144 YD-----------------PESPVTP------STPGSPPVSPGP------LSPGGTPGKH 174
               |                 | |:          |  ||  | |         :||   | |
Db       2664 LDFTHATGSEQVELDTCLLAEEPMQSLHREHGELPPEPPTLPQPELCAVDTAPGYVAGAH 2723

Qy        175 VCG--HHLHTVGGVVERDVCHRCR-----------------HK-------------RWH 201
               |   : | |  : : ||  | :                  |:                | |
Db       2724 QFGLSQNSHLVLPLNQSDVRKRLQVQLSIRTFASSGLIYYVAHQNQMDYATLQLQEGRLH 2783

Qy        202 FIKPTNKSRE--SRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLMSVSGAE----TVN 255
              |:    | |   | |       :|| |::   | |:     :|:: |:| | |    ||
Db       2784 FMFDLGKGRTKVSHP------ALLSDGKWHTVKTEY----IKRKAFMTVDGQESPSVTVV 2833

Qy        256 GEVPATPVKRE 266
              |       |:|:
Db       2834 GNATTLDVERK 2844


RESULT 14
GNVSPD
genome polyprotein - plum pox virus (strain D)
N;Contains: 29K protein; 34K protein; coat protein (capsid protein); cylindrical inclusic
C;Species: plum pox virus, PPV
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S06929; JA0078
R;Teycheney, P.Y.; Tavert, G.; Delbos, R.; Ravelonandro, M.; Dunez, J.
Nucleic Acids Res. 17, 10115-10116, 1989
A;Title: The complete nucleotide sequence of plum pox virus RNA (strain D).
A;Reference number: S06929; MUID:90098790; PMID:2602121
A;Accession: S06929
A;Status: translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-3141 <TEY>
A;Cross-references: UNIPROT:P13529; UNIPARC:UPI0000174B40; EMBL:X16415
R;Ravelonandro, M.; Varveri, C.; Delbos, R.; Dunez, J.
J. Gen. Virol. 69, 1509-1516, 1988
A;Title: Nucleotide sequence of the capsid protein gene of plum pox potyvirus.
A;Reference number: JA0078
A;Accession: JA0078
A;Molecule type: genomic RNA
A;Residues: 2709,'QT',2712,'L',2714,2810-3141 <RAV>
A;Cross-references: UNIPARC:UPI0000170FC8; GB:D00298; NID:g222410; PIDN:BAA00210.1; PID:g
A;Note: part of this sequence, including the amino end of the mature protein, was confirm
C;Superfamily: tobacco etch virus genome polyprotein
C;Keywords: ATP; coat protein; cylindrical inclusion protein; genome-linked protein; incl
F;1-301/Product: 34K protein #status predicted <PRO>
F;302-915/Product: helper component protein #status predicted <HCP>
```

```
F;916-1169/Product: 29K protein #status predicted <PRT>
F;1170-1857/Product: cylindrical inclusion protein #status predicted <CIP>
F;1254-1261/Region: nucleotide-binding motif A (P-loop)
F;1339-1344/Region: nucleotide-binding motif B
F;1343-1346/Region: DEXH motif
F;1858-2293/Product: nuclear inclusion protein a #status predicted <NIA>
F;2294-2811/Product: nuclear inclusion protein b #status predicted <NIB>
F;2812-3141/Product: coat protein #status predicted <COP>
F;1920/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

  Query Match             6.5%;  Score 92.5;  DB 1;  Length 3141;
  Best Local Similarity   22.0%;  Pred. No. 1e+02;
  Matches   41;  Conservative   30;  Mismatches   82;  Indels   33;  Gaps    7;

Qy         83 KKGYRCTTEAEQDIEEEKVEKIE-LNDSVNENSDTVGQIVHYIMKNEANAD-------VL 134
              || |  |  :| :||      :|   | :|:: ::   :||   : |   :       |:
Db       2775 KKLYTDTEASETEIE----RYLEAFYDDINDDGES-NVVVHQADEREDEEEVDAGKPIVV 2829

Qy        135 KAMVADNSLYDPESPVTPSTPGSPPV-----SPGPLSPGGTPGKHVCGHHLHTVGGVVER 189
              |  | : :  |   : |:   : |:      :|    |   | |  | | |
Db       2830 TAPAATSPILQPPPVIQPAPRTTAPMLNPIFTPATTQPATKPVSQVPGPQLQTFGTYGNE 2889

Qy        190 DVCHRCRHKRWHFIKPTNKSRESRPRRQGEVTVLSVGRFRVTKVEHKSNQKERRSLMSVS 249
              |              |:| :      |  :|   |:| | | :::  ::   : ||  |
Db       2890 DA-----------SPSNSNALVNTNRDRDVDAGSIGTFTVPRLKAMTS---KLSLPKVK 2934

Qy        250 GAETVN 255
              |   :|
Db       2935 GKAIMN 2940


RESULT 15
T29774
hypothetical protein T21D9.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29774
R;Nhan, M.; Hawkins, J.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid T21D9.
A;Reference number: Z20683
A;Accession: T29774
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-120 <NHA>
A;Cross-references: UNIPROT:Q22652; UNIPARC:UPI00000827D8; EMBL:U40937; PIDN:AAA81694.1;
C;Genetics:
A;Gene: CESP:T21D9.2
A;Introns: 56/2; 78/1
C;Superfamily: Caenorhabditis elegans hypothetical protein T21D9.2

  Query Match             6.5%;  Score 92;  DB 2;  Length 120;
  Best Local Similarity   26.5%;  Pred. No. 2;
  Matches   26;  Conservative   27;  Mismatches   29;  Indels   16;  Gaps    5;

Qy         30 DNGSSRTLHSRTETTPSPSNDTGNGHPEYIAYALVPVFFIMGL--FGVLICHLLKKKGYR 87
              || : |: :|  | || ||            | ::|:  :: |   |::|  :| ::  |
Db         22 DNIALRSTNSNDELTPVPST----------KYEILPIVMLISLSMSGIVIILILTRR--R 69

Qy         88 CTTEAEQDIEEEKVEKIELNDSVNENSDTVGQIVHYIM 125
              |  :|| |:    : :|::|| |::  :|  ::: :|:
Db         70 CLQDAESDL-TVALSEIDMND-VHQIFETTQEVLIFIL 105


Search completed: February 23, 2006, 12:08:34
Job time : 42 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 23, 2006, 15:09:53 ; Search time 1204 Seconds
(without alignments)
8866.916 Million cell updates/sec

Title: US-10-620-562-1
Perfect score: 1291
Sequence: 1 cccacgcgtccggcccgccg.....cttacgcaaaaaaaaaaaaa 1291

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1291	100.0	1291	3	US-09-910-562-1	Sequence 1, Appli
	2	1291	100.0	1291	9	US-10-620-562-1	Sequence 1, Appli
	3	1196.6	92.7	1603	3	US-09-933-767-56	Sequence 56, Appl
	4	1196.6	92.7	1603	5	US-10-004-860-56	Sequence 56, Appl
	5	1196.6	92.7	1603	5	US-10-023-282-56	Sequence 56, Appl
	6	226.2	17.5	591	4	US-09-925-065A-84619	Sequence 84619, A
	7	115.4	8.9	694	8	US-10-363-345A-5597	Sequence 5597, Ap
c	8	115.4	8.9	694	8	US-10-363-345A-5598	Sequence 5598, Ap
	9	115.4	8.9	694	9	US-10-363-483A-5597	Sequence 5597, Ap
c	10	115.4	8.9	694	9	US-10-363-483A-5598	Sequence 5598, Ap
c	11	104.6	8.1	694	8	US-10-363-345A-5599	Sequence 5599, Ap
	12	104.6	8.1	694	8	US-10-363-345A-5600	Sequence 5600, Ap
c	13	104.6	8.1	694	9	US-10-363-483A-5599	Sequence 5599, Ap
	14	104.6	8.1	694	9	US-10-363-483A-5600	Sequence 5600, Ap
	15	74	5.7	2205	4	US-09-925-065A-2551	Sequence 2551, Ap
	16	74	5.7	2205	4	US-09-925-065A-2552	Sequence 2552, Ap
	17	74	5.7	2205	4	US-09-925-065A-2553	Sequence 2553, Ap
	18	74	5.7	2480	4	US-09-925-065A-720502	Sequence 720502,
	19	74	5.7	2480	4	US-09-925-065A-720503	Sequence 720503,
	20	74	5.7	2480	4	US-09-925-065A-720504	Sequence 720504,
	21	74	5.7	2480	4	US-09-925-065A-720505	Sequence 720505,
	22	73	5.7	488	3	US-09-918-995-636	Sequence 636, App
	23	72.6	5.6	582	3	US-09-910-562-3	Sequence 3, Appli
	24	72.6	5.6	582	9	US-10-620-562-3	Sequence 3, Appli
	25	72.6	5.6	1793	6	US-10-120-988-2	Sequence 2, Appli
c	26	72.6	5.6	2886	3	US-09-764-875-82	Sequence 82, Appl
	27	58.6	4.5	1290	3	US-09-057-951-3	Sequence 3, Appli
	28	58.6	4.5	1290	5	US-10-105-150-3	Sequence 3, Appli
	29	58.6	4.5	2570	3	US-09-057-951-1	Sequence 1, Appli
	30	58.6	4.5	2570	5	US-10-105-150-1	Sequence 1, Appli
	31	58.6	4.5	2638	5	US-10-146-574-1	Sequence 1, Appli
	32	58.6	4.5	2703	3	US-09-836-607-1	Sequence 1, Appli
	33	58.6	4.5	2703	3	US-09-421-112-1	Sequence 1, Appli
	34	58.6	4.5	2732	5	US-10-157-031-264	Sequence 264, App
	35	58.6	4.5	3402	7	US-10-322-281-521	Sequence 521, App
	36	57	4.4	2606	9	US-10-956-157-2615	Sequence 2615, Ap
	37	50.6	3.9	2472	7	US-10-322-281-514	Sequence 514, App
	38	50.6	3.9	2479	5	US-10-146-574-3	Sequence 3, Appli
c	39	49.8	3.9	580	3	US-09-796-692-330	Sequence 330, App
c	40	49.8	3.9	580	3	US-09-796-692-9263	Sequence 9263, Ap
c	41	49.8	3.9	580	5	US-10-040-862-330	Sequence 330, App
c	42	49.8	3.9	580	5	US-10-040-862-9263	Sequence 9263, Ap
c	43	49.8	3.9	580	6	US-10-057-475B-330	Sequence 330, App
c	44	49.8	3.9	580	6	US-10-057-475B-9263	Sequence 9263, Ap
c	45	49.8	3.9	580	6	US-10-154-884B-330	Sequence 330, App

ALIGNMENTS

RESULT 1
US-09-910-562-1
; Sequence 1, Application US/09910562
; Patent No. US20020098163A1
; GENERAL INFORMATION:
; APPLICANT: Zeng et al.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptors TR21 and TR22
; FILE REFERENCE: PF530
; CURRENT APPLICATION NUMBER: US/09/910,562
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 60/221,143
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,116
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-910-562-1

Query Match 100.0%; Score 1291; DB 3; Length 1291;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
```

```
Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAA 360

Qy        361 CAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTG 420

Qy        421 TGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGA 480

Qy        481 ATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCG 540

Qy        541 TGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGA 600

Qy        601 CGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGG 660

Qy        661 ATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCA 720

Qy        721 GAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAG 780

Qy        781 TTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTG 840

Qy        841 GGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCA 900

Qy        901 CAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGTGAGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGTGAGG 960

Qy        961 AGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAGGGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAGGGAA 1020

Qy       1021 GTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTGTCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTGTCTG 1080

Qy       1081 CCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGTGTGA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGTGTGA 1140

Qy       1141 TAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAAACTA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAAACTA 1200

Qy       1201 GAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTAGTGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTAGTGT 1260

Qy       1261 TGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||||||||
Db       1261 TGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
```

RESULT 2
US-10-620-562-1
; Sequence 1, Application US/10620562
; Publication No. US20050090436A1

```
; GENERAL INFORMATION:
;  APPLICANT: Zeng, Z et al.
;  TITLE OF INVENTION: Human Tumor Necrosis Factor receptors TR21 and TR22
;  FILE REFERENCE: PF530C1
;  CURRENT APPLICATION NUMBER: US/10/620,562
;  CURRENT FILING DATE:  2003-07-17
;  PRIOR APPLICATION NUMBER: 09/910,562
;  PRIOR FILING DATE: 2001-07-23
;  PRIOR APPLICATION NUMBER: 60/221,143
;  PRIOR FILING DATE: 2000-07-27
;  PRIOR APPLICATION NUMBER: 60/220,116
;  PRIOR FILING DATE: 2000-07-24
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1291
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-10-620-562-1

  Query Match             100.0%;  Score 1291;  DB 9;  Length 1291;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1291;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAA 360

Qy        361 CAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTG 420

Qy        421 TGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGA 480

Qy        481 ATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCG 540

Qy        541 TGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGA 600

Qy        601 CGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGG 660

Qy        661 ATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCA 720

Qy        721 GAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAG 780
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAG 780

Qy        781 TTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTG 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTG 840

Qy        841 GGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCA 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCA 900

Qy        901 CAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGTGAGG 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGTGAGG 960

Qy        961 AGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAGGGAA 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAGGGAA 1020

Qy       1021 GTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTGTCTG 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTGTCTG 1080

Qy       1081 CCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGTGTGA 1140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGTGTGA 1140

Qy       1141 TAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAAACTA 1200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAAACTA 1200

Qy       1201 GAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTAGTGT 1260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTAGTGT 1260

Qy       1261 TGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
                 ||||||||||||||||||||||||||||||
Db       1261 TGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291


RESULT 3
US-09-933-767-56
; Sequence 56, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P2
;  CURRENT APPLICATION NUMBER: US/09/933,767
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: PCT/US01/05614
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: 60/184,836
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/193,170
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 09/205,258
;  PRIOR FILING DATE: 1998-12-04
;  PRIOR APPLICATION NUMBER: PCT/US98/11422
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/048,885
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,375
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,881
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,880
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,896
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,020
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,876
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,895
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,884
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,894
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,971
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,964
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,882
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,899
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,893
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,900
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,901
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,892
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,915
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,019
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,970
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,972
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,916
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,373
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,875
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/049,374
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,917
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,949
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,974
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,883
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,897
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,898
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,962
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,963
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,877
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/048,878
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/068,054
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,064
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,053
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/070,923
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/073,160
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/073,159
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/073,165
;  PRIOR FILING DATE: 1998-01-30
```

```
;  PRIOR APPLICATION NUMBER: 60/073,164
;  PRIOR FILING DATE: 1998-01-30
;  PRIOR APPLICATION NUMBER: 60/085,925
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,921
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,923
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/085,922
;  PRIOR FILING DATE: 1998-05-18
;  PRIOR APPLICATION NUMBER: 60/092,921
;  PRIOR FILING DATE: 1998-07-15
;  PRIOR APPLICATION NUMBER: 60/094,657
;  PRIOR FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1245
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
;   LENGTH: 1603
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (328)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (336)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (341)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (788)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-933-767-56

  Query Match             92.7%;  Score 1196.6;  DB 3;  Length 1603;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1268;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    6;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||||||  |  :      | ||||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||| | ||||||||||||||:|||:||||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||

Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296


RESULT 4
US-10-004-860-56
; Sequence 56, Application US/10004860
; Publication No. US20030065160A1
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/004,860
;  CURRENT FILING DATE:  2001-12-07
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
;   LENGTH: 1603
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: SITE
;   LOCATION: (328)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (336)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (341)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (788)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-004-860-56

  Query Match             92.7%;  Score 1196.6;  DB 5;  Length 1603;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1268;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    6;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||||  |  :    | ||||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||| | ||||||||||||||:|||:|||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296


RESULT 5
US-10-023-282-56
; Sequence 56, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/023,282
;  CURRENT FILING DATE: 2001-12-20
;  EARLIER APPLICATION NUMBER: 09/205,258
;  EARLIER FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
```

; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 1603
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (328)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (336)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (341)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (788)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-023-282-56

Query Match 92.7%; Score 1196.6; DB 5; Length 1603;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1268; Conservative 5; Mismatches 13; Indels 9; Gaps 6;

```
Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||||||  |  :    | ||||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||||| | ||||||||||||||:|||:||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              |||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296


RESULT 6
US-09-925-065A-84619
; Sequence 84619, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84619
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-84619

  Query Match            17.5%;  Score 226.2;  DB 4;  Length 591;
  Best Local Similarity   98.7%;  Pred. No. 1.2e-57;
  Matches  228;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        175 CTCCGGACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCA 234
              ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 CTGTAGACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCA 170

Qy        235 GCAACGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCT 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAACGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCT 230

Qy        295 TTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTT 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTT 290

Qy        355 GTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAG 405
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 GTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAG 341


RESULT 7
US-10-363-345A-5597
; Sequence 5597, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 5597
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 5597
US-10-363-345A-5597

  Query Match             8.9%;  Score 115.4;  DB 8;  Length 694;
  Best Local Similarity   81.2%;  Pred. No. 6.4e-24;
  Matches  134;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         16 CGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAG 75
              || ||| | |||||||| | ||  ||  ||||||||||||||||||   ||||||||||
Db        246 CGTCGGTTTCGGAGCGGTTTTGTTTTTTCGAGCGCGGGACGCGGCGTTTTGGGGGAGGAG 305

Qy         76 GGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTG 135
              |||||||||||||||||||| | |||||| | |  |||||| || |||  ||| || ||
Db        306 GGCGAAGCGACGCGGCGATGGTTTCGCGGGTATTTTCGGGGTTCGTCGTTTTAGTCGTTG 365

Qy        136 CTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||| |||||||||||||| |||||||||| || |||||| | |||
Db        366 TTGTTTTCGTGGGAGGCGTCGTGAGTTCGTCGTTGGTGGTTTCGG 410


RESULT 8
US-10-363-345A-5598/c
; Sequence 5598, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 5598
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 5598
US-10-363-345A-5598

  Query Match             8.9%;  Score 115.4;  DB 8;  Length 694;
  Best Local Similarity   81.2%;  Pred. No. 6.4e-24;
  Matches  134;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         16 CGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAG 75
              || ||| | |||||||| | ||  ||  ||||||||||||||||||   ||||||||||
```

```
Db        449 CGTCGGTTTCGGAGCGGTTTTGTTTTTTCGAGCGCGGGACGCGGCGTTTTGGGGGAGGAG 390

Qy         76 GGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTG 135
              |||||||||||||||||||||| | |||||| | |  |||||| || |||  ||| || ||
Db        389 GGCGAAGCGACGCGGCGATGGTTTCGCGGGTATTTTCGGGGTTCGTCGTTTTAGTCGTTG 330

Qy        136 CTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||| |||||||||||||| |||||||||| || |||||| | |||
Db        329 TTGTTTTCGTGGGAGGCGTCGTGAGTTCGTCGTTGGTGGTTTCGG 285


RESULT 9
US-10-363-483A-5597
; Sequence 5597, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 5597
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 5597
US-10-363-483A-5597

  Query Match             8.9%;  Score 115.4;  DB 9;  Length 694;
  Best Local Similarity   81.2%;  Pred. No. 6.4e-24;
  Matches  134;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         16 CGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAG 75
              || ||| | |||||||| | ||  ||  |||||||||||||||||||   ||||||||||
Db        246 CGTCGGTTTCGGAGCGGTTTTGTTTTTTCGAGCGCGGGACGCGGCGTTTTGGGGGAGGAG 305

Qy         76 GGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTG 135
              |||||||||||||||||||| | |||||| | |  |||||| || |||  ||| || ||
Db        306 GGCGAAGCGACGCGGCGATGGTTTCGCGGGTATTTTCGGGGTTCGTCGTTTTAGTCGTTG 365

Qy        136 CTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||| |||||||||||||| |||||||||| || |||||| | |||
Db        366 TTGTTTTCGTGGGAGGCGTCGTGAGTTCGTCGTTGGTGGTTTCGG 410


RESULT 10
US-10-363-483A-5598/c
; Sequence 5598, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 5598
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 5598
US-10-363-483A-5598

  Query Match             8.9%;  Score 115.4;  DB 9;  Length 694;
  Best Local Similarity   81.2%;  Pred. No. 6.4e-24;
  Matches  134;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         16 CGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAG 75
              || ||| | |||||||| | ||  ||  |||||||||||||||||||   ||||||||||
Db        449 CGTCGGTTTCGGAGCGGTTTTGTTTTTTCGAGCGCGGGACGCGGCGTTTTGGGGGAGGAG 390

Qy         76 GGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTG 135
              |||||||||||||||||||| | |||||| | |  |||||| || |||  ||| || ||
Db        389 GGCGAAGCGACGCGGCGATGGTTTCGCGGGTATTTTCGGGGTTCGTCGTTTTAGTCGTTG 330

Qy        136 CTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||| |||||||||||||| |||||||||| || |||||| | |||
Db        329 TTGTTTTCGTGGGAGGCGTCGTGAGTTCGTCGTTGGTGGTTTCGG 285


RESULT 11
US-10-363-345A-5599/c
; Sequence 5599, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 5599
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 5599
US-10-363-345A-5599

  Query Match             8.1%;  Score 104.6;  DB 8;  Length 694;
  Best Local Similarity   73.2%;  Pred. No. 1.2e-20;
  Matches  134;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          2 CCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 61
              |||| ||  |  ||||||| ||||| | || |||| |||||||| ||||  ||||| ||
Db        463 CCACTCGATCAACCCGCCGACTCCGAAACGACTCTACCTTCCCGAACGCGAAACGCGACG 404

Qy         62 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 121
              ||||      |  |  |||| |||||||  ||||  |||||||  |||||||   |||||
Db        403 CCCTAAAAAAAAAAAAACGAAACGACGCGACGATAACTCCGCGAACACTCCCGAAATCCGC 344

Qy        122 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGA 181
              ||||||| ||||| || |||||||    |  |||||| | |||||||||  |  ||||| |
Db        343 CGTCCTAACCGCTACTATCTTCGTAAAAAACGCCGTAAATTCGCCGCTAATAACTCCGAA 284

Qy        182 CAA 184
              |||
Db        283 TAA 281


RESULT 12
US-10-363-345A-5600
; Sequence 5600, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
```

```
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 5600
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 5600
US-10-363-345A-5600

  Query Match             8.1%;  Score 104.6;  DB 8;  Length 694;
  Best Local Similarity   73.2%;  Pred. No. 1.2e-20;
  Matches  134;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          2 CCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 61
              |||| ||  |  ||||||| ||||| | || |||| ||||||||| ||||  ||||| ||
Db        232 CCACTCGATCAACCCGCCGACTCCGAAACGACTCTACCTTCCCGAACGCGAAACGCGACG 291

Qy         62 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 121
              ||||     |  |   |||| ||||||| ||||  |||||||  |||||||   |||||
Db        292 CCCTAAAAAAAAAAAACGAAACGACGCGACGATAACTCCGCGAACACTCCCGAAATCCGC 351

Qy        122 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGA 181
              ||||||| ||||| || |||||||   |  |||||| | |||||||||  |  ||||| |
Db        352 CGTCCTAACCGCTACTATCTTCGTAAAAAACGCCGTAAATTCGCCGCTAATAACTCCGAA 411

Qy        182 CAA 184
              ||
Db        412 TAA 414


RESULT 13
US-10-363-483A-5599/c
; Sequence 5599, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 5599
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 5599
US-10-363-483A-5599

  Query Match             8.1%;  Score 104.6;  DB 9;  Length 694;
  Best Local Similarity   73.2%;  Pred. No. 1.2e-20;
  Matches  134;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          2 CCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 61
              |||| ||  |  ||||||| ||||| | || |||| ||||||||| ||||  ||||| ||
Db        463 CCACTCGATCAACCCGCCGACTCCGAAACGACTCTACCTTCCCGAACGCGAAACGCGACG 404

Qy         62 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 121
              ||||     |  |   |||| ||||||| ||||  |||||||  |||||||   |||||
Db        403 CCCTAAAAAAAAAAAACGAAACGACGCGACGATAACTCCGCGAACACTCCCGAAATCCGC 344

Qy        122 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGA 181
              ||||||| ||||| || |||||||   |  |||||| | |||||||||  |  ||||| |
Db        343 CGTCCTAACCGCTACTATCTTCGTAAAAAACGCCGTAAATTCGCCGCTAATAACTCCGAA 284

Qy        182 CAA 184
              ||
Db        283 TAA 281


RESULT 14
US-10-363-483A-5600
; Sequence 5600, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 5600
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 5600
US-10-363-483A-5600

  Query Match             8.1%;  Score 104.6;  DB 9;  Length 694;
  Best Local Similarity   73.2%;  Pred. No. 1.2e-20;
  Matches  134;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          2 CCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 61
              |||| ||  |  ||||||| ||||| | || |||| ||||||||| ||||  ||||| ||
Db        232 CCACTCGATCAACCCGCCGACTCCGAAACGACTCTACCTTCCCGAACGCGAAACGCGACG 291

Qy         62 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 121
              ||||     |  |   |||| ||||||| ||||  |||||||  |||||||   |||||
Db        292 CCCTAAAAAAAAAAAACGAAACGACGCGACGATAACTCCGCGAACACTCCCGAAATCCGC 351

Qy        122 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGA 181
              ||||||| ||||| || |||||||   |  |||||| | |||||||||  |  ||||| |
Db        352 CGTCCTAACCGCTACTATCTTCGTAAAAAACGCCGTAAATTCGCCGCTAATAACTCCGAA 411

Qy        182 CAA 184
              ||
Db        412 TAA 414


RESULT 15
US-09-925-065A-2551
; Sequence 2551, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2551
;   LENGTH: 2205
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-2551

  Query Match             5.7%;  Score 74;  DB 4;  Length 2205;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-11;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 218

Qy        464 CATGAAAAATGAAG 477
              ||||||||||||||
Db        219 CATGAAAAATGAAG 232


Search completed: February 23, 2006, 17:25:18
Job time : 1205 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 23, 2006, 15:08:27 ; Search time 442 Seconds
(without alignments)
6228.248 Million cell updates/sec

Title: US-10-620-562-1
Perfect score: 1291
Sequence: 1 cccacgcgtccggcccgccg.....cttacgcaaaaaaaaaaaaa 1291

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7209121 seqs, 1066183437 residues

Total number of hits satisfying chosen parameters: 14418242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	226.2	17.5	591	6	US-09-925-065A-84619	Sequence 84619, A
2	74	5.7	2205	6	US-09-925-065A-2551	Sequence 2551, Ap
3	74	5.7	2205	6	US-09-925-065A-2552	Sequence 2552, Ap
4	74	5.7	2205	6	US-09-925-065A-2553	Sequence 2553, Ap
5	74	5.7	2480	6	US-09-925-065A-720502	Sequence 720502,
6	74	5.7	2480	6	US-09-925-065A-720503	Sequence 720503,
7	74	5.7	2480	6	US-09-925-065A-720504	Sequence 720504,
8	74	5.7	2480	6	US-09-925-065A-720505	Sequence 720505,
9	58.6	4.5	2638	12	US-11-042-814-1	Sequence 1, Appli
10	50.6	3.9	2479	12	US-11-042-814-3	Sequence 3, Appli
11	46.6	3.6	1155	6	US-09-925-065A-38365	Sequence 38365, A
c 12	45	3.5	1731	12	US-11-052-554A-548	Sequence 548, App
13	42.6	3.3	1189	8	US-10-055-877-29	Sequence 29, Appl
c 14	41.6	3.2	811	6	US-09-925-065A-28354	Sequence 28354, A
c 15	41.6	3.2	811	6	US-09-925-065A-28355	Sequence 28355, A
c 16	40.8	3.2	370	6	US-09-925-065A-184526	Sequence 184526,
17	40	3.1	5075	12	US-11-000-688-1133	Sequence 1133, Ap
c 18	40	3.1	5099	8	US-10-821-234-193	Sequence 193, App
c 19	39.8	3.1	3370	8	US-10-972-053-3	Sequence 3, Appli
c 20	39.2	3.0	172147	12	US-11-112-908-22	Sequence 22, Appl
c 21	39.2	3.0	188682	12	US-11-112-908-23	Sequence 23, Appl
22	39	3.0	47460	12	US-11-124-368A-2877	Sequence 2877, Ap
c 23	38.8	3.0	30140	12	US-11-052-544-29	Sequence 29, Appl
c 24	38.8	3.0	38703	12	US-11-052-544-28	Sequence 28, Appl
25	38.4	3.0	587	6	US-09-925-065A-320868	Sequence 320868,
c 26	38	2.9	383	6	US-09-925-065A-560204	Sequence 560204,
27	38	2.9	452	6	US-09-925-065A-644775	Sequence 644775,
c 28	38	2.9	153376	12	US-11-121-086-5	Sequence 5, Appli
c 29	38	2.9	172543	12	US-11-121-086-6	Sequence 6, Appli
30	37.8	2.9	645179	8	US-10-995-561-13293	Sequence 13293, A
31	37.6	2.9	289	6	US-09-925-065A-499724	Sequence 499724,
32	37.6	2.9	622	6	US-09-925-065A-862767	Sequence 862767,
33	37.6	2.9	3500	12	US-11-085-775-1	Sequence 1, Appli
34	37.2	2.9	587	6	US-09-925-065A-320869	Sequence 320869,
c 35	36.8	2.9	3657	9	US-11-072-512-352	Sequence 352, App
c 36	36.6	2.8	607	6	US-09-925-065A-767166	Sequence 767166,
c 37	36.6	2.8	2458	8	US-10-947-249-2	Sequence 2, Appli
c 38	36.6	2.8	2458	8	US-10-947-249-3	Sequence 3, Appli
39	36.6	2.8	2825	12	US-11-136-527-192	Sequence 192, App
40	36.6	2.8	22970	7	US-10-467-033-25	Sequence 25, Appl
41	36.6	2.8	22970	7	US-10-467-033-26	Sequence 26, Appl
c 42	36.4	2.8	1754	12	US-11-136-527-2453	Sequence 2453, Ap
c 43	36.4	2.8	3167	6	US-09-925-065A-21764	Sequence 21764, A
44	36.2	2.8	602	6	US-09-925-065A-493497	Sequence 493497,
45	36.2	2.8	602	6	US-09-925-065A-493498	Sequence 493498,

ALIGNMENTS

RESULT 1
US-09-925-065A-84619
; Sequence 84619, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84619
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-84619

Query Match 17.5%; Score 226.2; DB 6; Length 591;
Best Local Similarity 98.7%; Pred. No. 5e-53;
Matches 228; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      175 CTCCGGACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCA 234
            ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      111 CTGTAGACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCA 170

Qy      235 GCAACGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCT 294
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      171 GCAACGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCT 230

Qy      295 TTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTT 354
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        231 TTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTT 290

Qy        355 GTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAG 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 GTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAG 341


RESULT 2
US-09-925-065A-2551
; Sequence 2551, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2551
;   LENGTH: 2205
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-2551

  Query Match             5.7%;  Score 74;  DB 6;  Length 2205;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-10;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 218

Qy        464 CATGAAAAATGAAG 477
              ||||||||||||||
Db        219 CATGAAAAATGAAG 232


RESULT 3
US-09-925-065A-2552
; Sequence 2552, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2552
;   LENGTH: 2205
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-2552

  Query Match             5.7%;  Score 74;  DB 6;  Length 2205;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-10;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 218

Qy        464 CATGAAAAATGAAG 477
              ||||||||||||||
Db        219 CATGAAAAATGAAG 232


RESULT 4
US-09-925-065A-2553
; Sequence 2553, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2553
;   LENGTH: 2205
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-2553

  Query Match             5.7%;  Score 74;  DB 6;  Length 2205;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-10;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 218

Qy        464 CATGAAAAATGAAG 477
              ||||||||||||||
Db        219 CATGAAAAATGAAG 232


RESULT 5
US-09-925-065A-720502
; Sequence 720502, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
```

```
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 720502
;   LENGTH: 2480
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-720502

  Query Match             5.7%;  Score 74;  DB 6;  Length 2480;
  Best Local Similarity   100.0%;  Pred. No. 3e-10;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 493

Qy        464 CATGAAAAATGAAG 477
              ||||||||||||||
Db        494 CATGAAAAATGAAG 507


RESULT 6
US-09-925-065A-720503
; Sequence 720503, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 720503
;   LENGTH: 2480
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-720503

  Query Match             5.7%;  Score 74;  DB 6;  Length 2480;
  Best Local Similarity   100.0%;  Pred. No. 3e-10;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 493

Qy        464 CATGAAAAATGAAG 477
              ||||||||||||||
Db        494 CATGAAAAATGAAG 507


RESULT 7
US-09-925-065A-720504
; Sequence 720504, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 720504
;   LENGTH: 2480
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-720504

  Query Match             5.7%;  Score 74;  DB 6;  Length 2480;
  Best Local Similarity   100.0%;  Pred. No. 3e-10;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 493

Qy        464 CATGAAAAATGAAG 477
              ||||||||||||||
Db        494 CATGAAAAATGAAG 507


RESULT 8
US-09-925-065A-720505
; Sequence 720505, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 720505
;   LENGTH: 2480
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-720505

  Query Match             5.7%;  Score 74;  DB 6;  Length 2480;
  Best Local Similarity   100.0%;  Pred. No. 3e-10;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        404 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 AGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACAT 493

Qy        464 CATGAAAAATGAAG 477
              ||||||||||||||
Db        494 CATGAAAAATGAAG 507


RESULT 9
US-11-042-814-1
; Sequence 1, Application US/11042814
; Publication No. US20060024267A1
; GENERAL INFORMATION:
;  APPLICANT: Jing, Shuqian
;  APPLICANT:  Welcher, Andrew A
;  APPLICANT:  Boedigheimer, Michael J
;  APPLICANT:  Shu, Junyan
;  APPLICANT:  Gary M. Fox
;  TITLE OF INVENTION: TNFr/OPG-LIKE MOLECULES AND USES THEREOF
;  FILE REFERENCE: 01017/36854
;  CURRENT APPLICATION NUMBER: US/11/042,814
;  CURRENT FILING DATE:  2005-01-25
;  PRIOR APPLICATION NUMBER: US/10/146,574
;  PRIOR FILING DATE: 2002-05-15
;  PRIOR APPLICATION NUMBER: US/09/724,037
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2638
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (195)..(1484)
US-11-042-814-1

  Query Match             4.5%;  Score 58.6;  DB 12;  Length 2638;
  Best Local Similarity   55.5%;  Pred. No. 6.3e-06;
  Matches  137;  Conservative    0;  Mismatches  104;  Indels    6;  Gaps    1;

Qy        275 CGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCT 334
              |||  | |||||||| ||||   |||||| || || ||| |||   | ||| |||| ||
Db        692 CGCCATCGTCCCTGTCTTCTGCCTCATGGGGCTGTTGGGCATCCTGGTGTGCAACCTCCT 751

Qy        335 TAAGAAGAAAGGCTATCGTTGTACAACA------GAAGCAGAGCAAGATATCGAAGAGGA 388
               |||  ||| ||||| |  || ||  |       || |  | ||  |     | ||  |
Db        752 CAAGCGGAAGGGCTACCACTGCACGGCGCACAAGGAGGTCGGGCCCGGCCCTGGAGGTGG 811

Qy        389 AAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCA 448
              |     || ||  |         |    | || | |||  ||    |||||  |||||
Db        812 AGGCAGTGGAATCAACCCTGCCTACCGGACTGAGGATGCCAATGAGGACACCATTGGGGT 871

Qy        449 AATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGC 508
              | || | ||  ||||   | ||  ||| ||||||||  | | |  ||| | ||| |
Db        872 CCTGGTGCGCTTGATCACAGAGAAGAAAGAGAATGCTGCGGCCCTGGAGGAGCTGCTGAA 931

Qy        509 AGATAAC 515
              |||  ||
Db        932 AGAGTAC 938


RESULT 10
US-11-042-814-3
; Sequence 3, Application US/11042814
; Publication No. US20060024267A1
; GENERAL INFORMATION:
;  APPLICANT: Jing, Shuqian
;  APPLICANT:  Welcher, Andrew A
;  APPLICANT:  Boedigheimer, Michael J
;  APPLICANT:  Shu, Junyan
;  APPLICANT:  Gary M. Fox
;  TITLE OF INVENTION: TNFr/OPG-LIKE MOLECULES AND USES THEREOF
;  FILE REFERENCE: 01017/36854
;  CURRENT APPLICATION NUMBER: US/11/042,814
;  CURRENT FILING DATE:  2005-01-25
;  PRIOR APPLICATION NUMBER: US/10/146,574
;  PRIOR FILING DATE: 2002-05-15
;  PRIOR APPLICATION NUMBER: US/09/724,037
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2479
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (91)..(1398)
US-11-042-814-3

  Query Match             3.9%;  Score 50.6;  DB 12;  Length 2479;
  Best Local Similarity   51.6%;  Pred. No. 0.001;
  Matches  144;  Conservative    0;  Mismatches  129;  Indels    6;  Gaps    1;

Qy        239 CGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTAT 298
              || | ||| | |          | | |||    |      ||  | || ||||| || | | |
Db        573 CGGTCCTGAGGAGACGGCTGCCCAGTATGCAGTGATTGCCATCGTTCCTGTCTTTTGTCT 632

Qy        299 CATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTAC 358
              |||||| ||  | ||| |||  | ||| |||||||| |||  ||| ||||| | ||| ||
Db        633 CATGGGGCTTCTGGGCATCCTGGTGTGCAACCTGCTCAAGCGGAAGGGCTACCATTGCAC 692

Qy        359 AACA------GAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAA 412
              | |          ||||  | ||      |  |  | |      |  |  |       |
Db        693 AGCCCAAAAGGAAGTTGGGCCCAGCCCTGGTGGAGGAGGCAGCGGGATTAATCCTGCCTA 752

Qy        413 TGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAA 472
              |   | ||   |||  |||   |||||  ||||      | || | |   ||||   | ||
Db        753 TAGGACTGAAGATGCCAACGAGGACACCATTGGAGTCCTGGTGCGCCTGATCACAGAGAA 812

Qy        473 TGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGA 511
              ||| ||||||  |  | | |  ||| | || |   |||
Db        813 GAAAGAGAATGCAGCGGCCCTGGAGGAGCTGTTGAAAGA 851


RESULT 11
US-09-925-065A-38365
; Sequence 38365, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 38365
;   LENGTH: 1155
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-38365
```

```
  Query Match             3.6%;  Score 46.6;  DB 6;  Length 1155;
  Best Local Similarity   52.8%;  Pred. No. 0.01;
  Matches  123;  Conservative    0;  Mismatches  109;  Indels    1;  Gaps    1;

Qy       1050 CTGAGGGGGACTGAACATGATTACTTGTCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAA 1109
             ||  |||  | || | | |||   || | ||  | |   | |||   |  |||||||||
Db        185 CTTCAGGGAGCAGAGCTTCATTTTATGGCAGCTGAAACAGTATTGAGTAAGAGTCACAAA 244

Qy       1110 CT-TAGTGCCTCCAGGGGCTTGGCCTGTGTGATAATGAGGATAGAGGATTACTTGTGAGG 1168
              |  |||||||||||||  |||  ||   |  |   | |      ||  |   |||
Db        245 ATCAAATGCCTCCAGGGGCCAGGCAGGTACTACTAATTAGTAAAGCAGGTAGGTACAAGG 304

Qy       1169 CAATGTGGCATGGTGGGGATTGTGGCAAACTAGAATTCACATCACCCACCATATAGGGCT 1228
              |||   ||  |||||| || |||||||||| |||   ||| |  |   |  |   |
Db        305 CAAGAGGGAGTGGTGGTGACTGTGGCAAACCAGAGAGCACCTGCCTGGTCTGTTGATGGC 364

Qy       1229 TGCATTACCACGAGGCAGAAAGCACCTAGTGTTGCTGCATCTTCTTACGCAAA 1281
              ||| |  || | |   |  | | |   ||   |   |  | ||||||   ||
Db        365 TGCCTCTCCTCCATTTAATAGGTAGGCAGGTATATGGGCTAATCTTACCAGAA 417


RESULT 12
US-11-052-554A-548/c
; Sequence 548, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 548
;   LENGTH: 1731
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-548

  Query Match             3.5%;  Score 45;  DB 12;  Length 1731;
  Best Local Similarity   53.0%;  Pred. No. 0.033;
  Matches   96;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              |||| ||  |||||||  | |   | || | || | |||  ||||  |||   |  | |
Db        509 CCCAGGCCGCCGGCCCCGCCGTGGCCGATCAGCCCGGCCGCCCCGCCCGCACCGCCGGCC 450

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
                | || || |  || |    ||  || ||  | ||    | ||  | | | ||  |||
Db        449 TGACCGGTGGCGCCGGACCCGCCGTTGCCGCCGTTGCCGTACAGGATCCCGCCGGCCCCG 390

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||| |||  ||| | |      |   | |||||||||      | || || | | | |
Db        389 CCGGCCTGCCCGGTCCCCGGCGCCCCGTCGGCGCCGTGGCCGATCAGCGGGCGCCCCAGC 330

Qy        181 A 181
              |
Db        329 A 329


RESULT 13
US-10-055-877-29
; Sequence 29, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
;  APPLICANT: DeCristofaro, Marc
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Zhong, Mei
;  APPLICANT:  Anderson, David
;  APPLICANT:  Ballinger, Robert
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Ratelli, Luca
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Andrew, David
;  APPLICANT:  Mezes, Peter
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Burgess, Cahterine
;  APPLICANT:  Eisen, Andrew
;  APPLICANT:  Wolenc, Adam
;  APPLICANT:  Baumgartner, Jason
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Gusev, Vladimir
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Taupier Jr., Raymond
;  APPLICANT:  Pena, Carol
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Li, Li
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Boldog, Ference
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
;  FILE REFERENCE: 21402-251
;  CURRENT APPLICATION NUMBER: US/10/055,877
;  CURRENT FILING DATE:  2002-01-22
;  PRIOR APPLICATION NUMBER: 60/262,892
;  PRIOR FILING DATE: 2001-01-19
;  PRIOR APPLICATION NUMBER: 60/263,598
;  PRIOR FILING DATE: 2001-01-23
;  PRIOR APPLICATION NUMBER: 60/263,799
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: 60/264,117
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/264,139
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/264,478
;  PRIOR FILING DATE: 2001-01-26
;  PRIOR APPLICATION NUMBER: 60/263,351
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: 60/272,870
;  PRIOR FILING DATE: 2001-03-02
;  PRIOR APPLICATION NUMBER: 60/275,990
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/275,927
;  PRIOR FILING DATE: 2001-03-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 512
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
;   LENGTH: 1189
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (85)..(1143)
US-10-055-877-29

  Query Match             3.3%;  Score 42.6;  DB 8;  Length 1189;
  Best Local Similarity   48.2%;  Pred. No. 0.13;
  Matches  120;  Conservative    0;  Mismatches  129;  Indels    0;  Gaps    0;

Qy          9 TCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGG 68
              |||  |  |||  |  ||    | |||||  |||   |  || ||  |  ||   ||
Db        262 TCCTCCGTGCCCTCCTCGCCCAGCTTCTGCGCGCCCAGCCCGGGCACCGGCGGCGGCGGC 321

Qy         69 GGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTA 128
              || |  ||| |  ||| ||   ||        |  |||     ||| ||| |  |  |
Db        322 GGCGCGGGGGGCGGCGGCGGCTCGTCTCAGGCCGGGGGCGCCCCCGGGCCGCCGAGCGGG 381
```

```
Qy        129 GCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGG 188
              | | | |  | | ||| |||   | | | ||   || |||||| || || | ||     |
Db        382 GGCCCCGGCGCCGTCGGGGGCACCTCGGGGAAGCCGGCGCTGGAGGATCTGTACTGGATG 441

Qy        189 AGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGG 248
              ||| ||  |      ||| ||   | ||| ||  |     | |   | | ||  | | |
Db        442 AGCGGCTACCAGCATCACCTCAACCCCGAGGCGCTCAACCTGACGCCCGAGGACGCGGTG 501

Qy        249 AATGGACAC 257
              | |  | |
Db        502 GAGGCGCTC 510


RESULT 14
US-09-925-065A-28354/c
; Sequence 28354, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28354
;   LENGTH: 811
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-28354

  Query Match             3.2%;  Score 41.6;  DB 6;  Length 811;
  Best Local Similarity   56.6%;  Pred. No. 0.22;
  Matches   77;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy         43 CCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGC 102
              ||| | ||| | ||||| || | || | | |  ||     ||  |  |||  || ||||
Db        281 CCGGGGGCGTGGCGCGGGGCGCAGGTGCACGCAGGTTGGCCGCGGGCGCGCGGGGTCCGG 222

Qy        103 GGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTT 162
              || | |||||||||    ||   |  | || |  | | ||   |  ||||  | ||| |
Db        221 CGGGAGTCCCGGGGTGGTCCCCGCGTGGGGCCGGGGGCGTCCGCGCGGGCGGGGCGAGGT 162

Qy        163 CGCCGCTGGTGGCTCC 178
              ||  |  ||  ||| |
Db        161 CGGGGAGGGCAGCTGC 146


RESULT 15
US-09-925-065A-28355/c
; Sequence 28355, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28355
;   LENGTH: 811
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-28355

  Query Match             3.2%;  Score 41.6;  DB 6;  Length 811;
  Best Local Similarity   56.6%;  Pred. No. 0.22;
  Matches   77;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy         43 CCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGC 102
              ||| | ||| | ||||| || | || | | |  ||     ||  |  |||  || ||||
Db        281 CCGGGGGCGTGGCGCGGGGCGCAGGTGCACGCAGGTTGGCCGCGGGCGCGCGGGGTCCGG 222

Qy        103 GGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTT 162
              || | |||||||||    ||   |  | || |  | | ||   |  ||||  | ||| |
Db        221 CGGGAGTCCCGGGGTGGTCCCCGCGTGGGGCCGGGGGCGTCCGCGCGGGCGGGGCGAGGT 162

Qy        163 CGCCGCTGGTGGCTCC 178
              ||  |  ||  ||| |
Db        161 CGGGGAGGGCAGCTGC 146


Search completed: February 23, 2006, 15:20:22
Job time : 443 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 23, 2006, 14:49:57 ; Search time 6919 Seconds
(without alignments)
10606.286 Million cell updates/sec

Title: US-10-620-562-1
Perfect score: 1291
Sequence: 1 cccacgcgtccggcccgccg.....cttacgcaaaaaaaaaaaaa 1291

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1196.6	92.7	1603	6	BD192961	BD192961 207 human
	2	1196.6	92.7	1603	6	CQ821957	CQ821957 Sequence
	3	900	69.7	1320	6	BD123581	BD123581 Secretory
	4	900	69.7	1320	6	CS051293	CS051293 Sequence
	5	900	69.7	1320	6	AX136257	AX136257 Sequence
	6	900	69.7	1320	8	AK075468	AK075468 Homo sapi
	7	899	69.6	1364	8	BC039540	BC039540 Homo sapi
	8	685.2	53.1	726	6	BD123784	BD123784 Secretory
	9	685.2	53.1	726	6	CS051580	CS051580 Sequence
	10	685.2	53.1	726	6	AX136544	AX136544 Sequence
	11	680	52.7	680	6	CQ727837	CQ727837 Sequence
	12	622.2	48.2	2581	9	BC031198	BC031198 Mus muscu
	13	620.6	48.1	2481	9	BC066137	BC066137 Mus muscu
	14	620.6	48.1	2481	9	BC066160	BC066160 Mus muscu
c	15	519.6	40.2	161070	8	AC027607	AC027607 Homo sapi
	16	519.6	40.2	171279	14	AC036185	AC036185 Homo sapi
	17	355.2	27.5	3672	5	AJ851746	AJ851746 Gallus ga
	18	353.2	27.4	1166	5	CR523601	CR523601 Gallus ga
	19	311.4	24.1	2544	9	BC024679	BC024679 Mus muscu
c	20	239.4	18.5	171279	14	AC036185	AC036185 Homo sapi
c	21	219.6	17.0	264263	14	AC161946	AC161946 Bos tauru
c	22	178.2	13.8	197543	9	AC124129	AC124129 Mus muscu
	23	178.2	13.8	248595	14	AC158789	AC158789 Mus muscu
	24	174.2	13.5	1383	5	BC056115	BC056115 Xenopus l
	25	172.8	13.4	230845	14	AC098415	AC098415 Rattus no
c	26	172.6	13.4	162609	8	AC108022	AC108022 Homo sapi
	27	172.4	13.4	1295	5	BC070680	BC070680 Xenopus l
	28	145.2	11.2	2132	5	BC088389	BC088389 Danio rer
c	29	112	8.7	70892	14	AC027533	AC027533 Homo sapi
	30	111	8.6	181070	14	AC119888	AC119888 Mus muscu
	31	81.2	6.3	1764	9	BC089946	BC089946 Rattus no
	32	78	6.0	2090	9	BC038500	BC038500 Mus muscu
	33	75.6	5.9	1881	5	BC075421	BC075421 Xenopus t
c	34	75.4	5.8	157691	5	AL935143	AL935143 Zebrafish
	35	72.6	5.6	827	6	BD125145	BD125145 Primer fo
	36	72.6	5.6	827	6	BD126407	BD126407 Primer fo
	37	72.6	5.6	827	6	CQ780436	CQ780436 Sequence
	38	72.6	5.6	827	6	CQ781698	CQ781698 Sequence
	39	72.6	5.6	1086	6	CQ730722	CQ730722 Sequence
	40	72.6	5.6	1409	6	BD127456	BD127456 Primer fo
	41	72.6	5.6	1409	6	CQ783064	CQ783064 Sequence
	42	72.6	5.6	1409	8	AK075064	AK075064 Homo sapi
	43	72.6	5.6	1666	8	BC063469	BC063469 Homo sapi
	44	72.6	5.6	1793	6	AR541754	AR541754 Sequence
	45	72.6	5.6	2535	8	AY358163	AY358163 Homo sapi

ALIGNMENTS

RESULT 1
BD192961
LOCUS BD192961 1603 bp DNA linear PAT 17-JUL-2003
DEFINITION 207 human secreted proteins.
ACCESSION BD192961
VERSION BD192961.1 GI:33002700
KEYWORDS JP 2002516573-A/55.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1603)
AUTHORS Young,P., Greene,J.M., Ferrie,A.M., Ruben,S.M., Rosen,C.A., Hu,J.S., Olsen,H.S., Ebner,R., Brewer,L.A., Moore,P.A., Shi,Y., Florence,C., Florence,K., Lafleur,D.W., Ni,J., Fan,P., Wei,Y.F., Fischer,C.L., Soppet,D.R., Li,Y., Ze,Z., Kyaw,H., Yu,G.L., Feng,P., Dillon,P.J., Endress,G.A. and Carter,K.C.
TITLE 207 human secreted proteins
JOURNAL Patent: JP 2002516573-A 55 04-JUN-2002;
HUMAN GENOME SCIENCES INC
COMMENT PN JP 2002516573-A/55
PD 04-JUN-2002
PF 04-JUN-1998 JP 1999502815
PR 06-JUN-1997 US 60/048915,06-JUN-1997 US 60/048882 PR
06-JUN-1997 US 60/048892,06-JUN-1997 US 60/048901 PR
06-JUN-1997 US 60/048900,06-JUN-1997 US 60/048893 PR
06-JUN-1997 US 60/048964,06-JUN-1997 US 60/048884 PR
06-JUN-1997 US 60/048894,06-JUN-1997 US 60/048971 PR
06-JUN-1997 US 60/048885,06-JUN-1997 US 60/049375 PR
06-JUN-1997 US 60/048881,06-JUN-1997 US 60/048880 PR
06-JUN-1997 US 60/048896,06-JUN-1997 US 60/049020 PR
06-JUN-1997 US 60/048876,06-JUN-1997 US 60/048895 PR
06-JUN-1997 US 60/049019,06-JUN-1997 US 60/048916 PR
06-JUN-1997 US 60/048970,06-JUN-1997 US 60/048972 PR
06-JUN-1997 US 60/048949,06-JUN-1997 US 60/048974 PR
06-JUN-1997 US 60/048883,06-JUN-1997 US 60/048897 PR
06-JUN-1997 US 60/048898,06-JUN-1997 US 60/049373 PR
06-JUN-1997 US 60/048917,06-JUN-1997 US 60/048962 PR
06-JUN-1997 US 60/048878,06-JUN-1997 US 60/049374 PR
06-JUN-1997 US 60/048875,06-JUN-1997 US 60/048899 PR
06-JUN-1997 US 60/048877,06-JUN-1997 US 60/048963 PR
05-SEP-1997 US 60/057651,05-SEP-1997 US 60/057769 PR

```
            05-SEP-1997 US    60/057643,05-SEP-1997 US    60/057645 PR
            05-SEP-1997 US    60/057668,05-SEP-1997 US    60/057635 PR
            05-SEP-1997 US    60/057627,05-SEP-1997 US    60/057667 PR
            05-SEP-1997 US    60/057666,05-SEP-1997 US    60/057764 PR
            05-SEP-1997 US    60/057644,05-SEP-1997 US    60/057765 PR
            05-SEP-1997 US    60/057762,05-SEP-1997 US    60/057775 PR
            05-SEP-1997 US    60/057634,05-SEP-1997 US    60/057777 PR
            05-SEP-1997 US    60/057628,05-SEP-1997 US    60/057776 PR
            05-SEP-1997 US    60/057760,05-SEP-1997 US    60/057761 PR
            05-SEP-1997 US    60/057771,05-SEP-1997 US    60/057770 PR
            05-SEP-1997 US    60/057649,05-SEP-1997 US    60/057774 PR
            05-SEP-1997 US    60/057648,05-SEP-1997 US    60/057642 PR
            05-SEP-1997 US    60/057629,05-SEP-1997 US    60/057778 PR
            05-SEP-1997 US    60/057763,05-SEP-1997 US    60/057584 PR
            05-SEP-1997 US    60/057654,05-SEP-1997 US    60/057646 PR
            05-SEP-1997 US    60/057662,05-SEP-1997 US    60/057650 PR
            05-SEP-1997 US    60/057661,05-SEP-1997 US    60/057647 PR
            18-DEC-1997 US    60/070923
            PI   PAUL YOUNG,JOHN M GREENE,ANN M FERRIE,STEVEN M RUBEN,CRAIG A
            PI   ROSEN,
            PI   JING SHAN HU,HENRIK S OLSEN,REINHARD EBNER,LAURIE A BREWER, PI
              PAUL A MOORE,
            PI   YANGGU SHI,CHARLES FLORENCE,KIMBERLY FLORENCE,DAVID W LAFLEUR,
            PI   JIAN NI,
            PI   PING FAN,YING FEI WEI,CARRIE L FISCHER,DANIEL R SOPPET,YI LI,
            PI   ZHIZHEN ZENG,
            PI   HLA KYAW,GUO LIANG YU,PING FENG,PATRICK J DILLON,GREGORY A  PI
              ENDRESS,
            PI   KENNETH C CARTER
            PC   A01N37/18,A01N43/04,C12Q1/00,C12Q1/02,C12Q1/68,C12N5/00 PC
            ,C12N5/06,C12N15/00,
            PC   C12N15/06,C12N15/09,C12N15/10,C12N15/11,G01N33/53 CC
            Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES                Location/Qualifiers
     source             1. .1603
                        /organism="synthetic construct"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32630"
ORIGIN

  Query Match             92.7%;  Score 1196.6;  DB 6;  Length 1603;
  Best Local Similarity   97.9%;  Pred. No. 1.8e-276;
  Matches 1268;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    6;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||| |  :     | ||||||||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||||| | |||||||||||||:|||:||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296


RESULT 2
CQ821957
LOCUS       CQ821957               1603 bp    DNA     linear   PAT 21-JUN-2004
DEFINITION  Sequence 56 from Patent EP1428833.
ACCESSION   CQ821957
VERSION     CQ821957.1  GI:49020602
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
          Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P., Greene,J.M., Ferrie,A.M., Ruben,S.M., Rosen,C.A.,
            Hu,J.S., Olsen,H.S., Ebner,R., Brewer,L.A., More,P.A., Shi,Y.,
            Florence,C., Florence,K., LaFleur,D.W., Ni,J., Fan,P., Wei,Y.F.,
            Fischer,C.L., Soppet,D.R., Li,Y., Zeng,Z., Kyaw,H., Yu,G.L.,
            Feng,P., Dillon,P.J., Endress,G.A. and Carter,K.C.
  TITLE     207 human secreted proteins
  JOURNAL   Patent: EP 1428833-A 56 16-JUN-2004;
            Human Genome Sciences, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1603
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    328
                     /note="n equals a,t,g, or c"
     misc_feature    336
                     /note="n equals a,t,g, or c"
     misc_feature    341
                     /note="n equals a,t,g, or c"
     misc_feature    788
                     /note="n equals a,t,g, or c"
ORIGIN

  Query Match             92.7%;  Score 1196.6;  DB 6;  Length 1603;
  Best Local Similarity   97.9%;  Pred. No. 1.8e-276;
  Matches 1268;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    6;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||||| |  :     | ||||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||| | ||||||||||||:|||:||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296


RESULT 3
BD123581
LOCUS       BD123581                1320 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Secretory protein or membrane protein.
ACCESSION   BD123581
VERSION     BD123581.1  GI:23218526
KEYWORDS    JP 2002017376-A/90.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1320)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
            Hayashi,K.
  TITLE     Secretory protein or membrane protein
  JOURNAL   Patent: JP 2002017376-A 90 22-JAN-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002017376-A/90
            PD   22-JAN-2002
            PF   07-JUL-2000 JP 2000253173
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,YURI KAWAI,TOMOYASU
            PI   SUGIYAMA,
            PI   KOJI HAYASHI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
```

```
           PC   C12P21/02,C12Q1/68//C12P21/08,C12N15/00,C12N5/00 CC
           Secretory protein or membrane protein
           FH   Key             Location/Qualifiers
           FT   CDS             (90). .(902).
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            69.7%;  Score 900;  DB 6;  Length 1320;
  Best Local Similarity  99.4%;  Pred. No. 3.2e-205;
  Matches  903;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 480

Qy        484 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 540

Qy        544 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 600

Qy        604 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 660

Qy        664 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 720

Qy        724 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 780

Qy        784 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 840

Qy        844 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 900

Qy        904 AGTAGCAG 911
              ||||||||
Db        901 AGTAGCAG 908


RESULT 4
CS051293
LOCUS       CS051293               1320 bp    DNA     linear   PAT 23-MAR-2005
DEFINITION  Sequence 179 from Patent EP1514933.
ACCESSION   CS051293
VERSION     CS051293.1  GI:61890472
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
            Hayashi,K.
  TITLE     Secretory protein or membrane protein
  JOURNAL   Patent: EP 1514933-A 179 16-MAR-2005;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             90. .905
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAI72141.1"
                     /db_xref="GI:61890473"
                     /translation="MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRALHSRTETT
                     PSPSNDTGNGHPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVE
                     KIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPV
                     SPGPLSPGGTPGKHVCGHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQG
                     EVTVLSVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGTE"
ORIGIN

  Query Match            69.7%;  Score 900;  DB 6;  Length 1320;
  Best Local Similarity  99.4%;  Pred. No. 3.2e-205;
  Matches  903;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420
```

```
Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 480

Qy        484 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 540

Qy        544 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 600

Qy        604 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 660

Qy        664 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 720

Qy        724 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 780

Qy        784 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 840

Qy        844 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 900

Qy        904 AGTAGCAG 911
              ||||||||
Db        901 AGTAGCAG 908
```

```
RESULT 5
AX136257
LOCUS       AX136257               1320 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 179 from Patent EP1067182.
ACCESSION   AX136257
VERSION     AX136257.1  GI:14272663
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
            Hayashi,K.
  TITLE     Secretory protein or membrane protein
  JOURNAL   Patent: EP 1067182-A 179 10-JAN-2001;
            Helix Research Institute (JP)
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             90. .905
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC39770.1"
                     /db_xref="GI:14272664"
                     /translation="MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRALHSRTETT
                     PSPSNDTGNGHPEYIAYALVPVFFIMGLPGVLICHLLKKKGYRCTTEAEQDIEEEKVE
                     KIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPV
                     SPGPLSPGGTPGKHVCGHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQG
                     EVTVLSVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGTE"
ORIGIN
```

```
  Query Match           69.7%;  Score 900;  DB 6;  Length 1320;
  Best Local Similarity   99.4%;  Pred. No. 3.2e-205;
  Matches  903;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 480

Qy        484 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 540

Qy        544 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 600

Qy        604 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 660

Qy        664 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 720

Qy        724 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 780

Qy        784 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 840

Qy        844 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 900

Qy        904 AGTAGCAG 911
              ||||||||
Db        901 AGTAGCAG 908
```

```
RESULT 6
AK075468
LOCUS       AK075468               1320 bp    mRNA    linear   PRI 09-JUL-2005
DEFINITION  Homo sapiens cDNA PSEC0162 fis, clone PLACE1009772.
```

```
ACCESSION   AK075468
VERSION     AK075468.1  GI:22761572
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Otsuki,T., Ota,T., Nishikawa,T., Hayashi,K., Suzuki,Y.,
            Yamamoto,J., Wakamatsu,A., Kimura,K., Sakamoto,K., Hatano,N.,
            Kawai,Y., Ishii,S., Saito,K., Kojima,S., Sugiyama,T., Ono,T.,
            Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A.,
            Okumura,K., Nagai,K., Sugano,S. and Isogai,T.
  TITLE     Signal Sequence and Keyword Trap in silico for Selection of
            Full-Length Human cDNAs Encoding Secretion or Membrane Proteins
            from Oligo-Capped cDNA Libraries
  JOURNAL   DNA Res. 12, 117-126 (2005)
REFERENCE   2
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Kawai-Hio,Y., Hayashi,K.,
            Ishii,S., Saito,K., Yamamoto,J., Wakamatsu,A., Nagai,T.,
            Nakamura,Y., Nagahari,K., Sugano,S. and Isogai,T.
  TITLE     HRI human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1320)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     HRI human cDNA sequencing project; cDNA 5'- & 3'-end one pass
            sequencing, clone selection and full insert sequencing: Helix
            Research Institute (supported by Japan Key Technology Center etc.);
            cDNA library construction: Institute of Medical Science, University
            of Tokyo, Laboratory of Genome Structure, Human Genome Center.
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1009772"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="cloning vector: pME18SFL3"
     CDS             90. .905
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC11638.1"
                     /db_xref="GI:22761573"
                     /translation="MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRALHSRTETT
                     PSPSNDTGNGHPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVE
                     KIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPV
                     SPGPLSPGGTPGKHVCGHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQG
                     EVTVLSVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGTE"
ORIGIN

  Query Match           69.7%;  Score 900;  DB 8;  Length 1320;
  Best Local Similarity  99.4%;  Pred. No. 3.2e-205;
  Matches  903;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 480

Qy        484 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 540

Qy        544 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 600

Qy        604 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 660

Qy        664 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 720

Qy        724 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 780

Qy        784 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 840

Qy        844 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 900

Qy        904 AGTAGCAG 911
              ||||||||
Db        901 AGTAGCAG 908


RESULT 7
BC039540
LOCUS       BC039540               1364 bp    mRNA    linear   PRI 07-OCT-2003
DEFINITION  Homo sapiens mRNA similar to expressed sequence AA536743 (cDNA
            clone MGC:50583 IMAGE:5748847), complete cds.
ACCESSION   BC039540
VERSION     BC039540.1  GI:24660026
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1364)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
```

```
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1364)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 88 Row: n Column: 22.
FEATURES              Location/Qualifiers
     source           1. .1364
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="MGC:50583 IMAGE:5748847"
                      /tissue_type="Pancreas, Spleen, adult pooled"
                      /clone_lib="NIH_MGC_120"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     CDS              89. .904
                      /codon_start=1
                      /product="Similar to expressed sequence AA536743"
                      /protein_id="AAH39540.1"
                      /db_xref="GI:24660027"
                      /translation="MAPRALPGSAVLAAAVFVGGAVSSPLVAPDNGSSRTLHSRTETT
                      PSPSNDTGNGHPEYIAYALVPVFFIMGLFGVLICHLLKKKGYRCTTEAEQDIEEEKVE
                      KIELNDSVNENSDTVGQIVHYIMKNEANADVLKAMVADNSLYDPESPVTPSTPGSPPV
                      SPGPLSPGGTPGKHVCGHHLHTVGGVVERDVCHRCRHKRWHFIKPTNKSRESRPRRQG
                      EVTVLSVGRFRVTKVEHKSNQKERRSLMSVSGAETVNGEVPATPVKRERSGTE"
ORIGIN

  Query Match             69.6%;  Score 899;  DB 8;  Length 1364;
  Best Local Similarity   99.4%;  Pred. No. 5.7e-205;
  Matches  902;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 CGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCC 64
              | ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCC 60

Qy         65 TGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGT 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGT 120

Qy        125 CCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAA 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAA 180

Qy        185 TGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATAC 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATAC 240

Qy        245 TGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGG 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGG 300

Qy        305 TCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGA 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGA 360

Qy        365 AGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAA 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAA 420

Qy        425 TGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGC 484
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        421 TGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCAAATGC 480

Qy        485 TGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGAC 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGAC 540

Qy        545 CCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCC 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCC 600

Qy        605 AGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGT 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGT 660

Qy        665 GTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGA 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGA 720

Qy        725 GAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTAC 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTAC 780

Qy        785 AAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGC 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGC 840

Qy        845 TGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAGA 900

Qy        905 GTAGCAG 911
              |||||||
Db        901 GTAGCAG 907


RESULT 8
BD123784
LOCUS       BD123784                726 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Secretory protein or membrane protein.
ACCESSION   BD123784
VERSION     BD123784.1  GI:23218729
KEYWORDS    JP 2002017376-A/293.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
            Hayashi,K.
  TITLE     Secretory protein or membrane protein
  JOURNAL   Patent: JP 2002017376-A 293 22-JAN-2002;
```

```
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002017376-A/293
            PD   22-JAN-2002
            PF   07-JUL-2000 JP 2000253173
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,YURI KAWAI,TOMOYASU
            PI   SUGIYAMA,
            PI   KOJI HAYASHI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,C12N15/00,C12N5/00 CC
            Secretory protein or membrane protein
            FH   Key             Location/Qualifiers
            FT   source          1. .726
            FT                   /organism='Homo sapiens (human)'.
FEATURES              Location/Qualifiers
     source           1. .726
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match            53.1%;  Score 685.2;  DB 6;  Length 726;
  Best Local Similarity  98.3%;  Pred. No. 1.2e-153;
  Matches  712;  Conservative    0;  Mismatches   10;  Indels    2;  Gaps    2;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||   | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCG-AAT 482
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAAAT 480

Qy        483 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 540

Qy        543 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 600

Qy        603 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGAT 662
              |||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||
Db        601 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCAAAAGGGAT 660

Qy        663 GTGTGTCAT-CGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAG 721
              ||||||||| ||||||| |||||| |||||||||||||||||||||||||||| ||||||
Db        661 GTGTGTCATCCGGTGTAAGCACAANCGGTGGCACTTTATAAAGCCCACTAACANGTCCAG 720

Qy        722 AGAG 725
              ||||
Db        721 AGAG 724


RESULT 9
CS051580
LOCUS       CS051580                726 bp    DNA     linear   PAT 23-MAR-2005
DEFINITION  Sequence 466 from Patent EP1514933.
ACCESSION   CS051580
VERSION     CS051580.1  GI:61890757
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
            Hayashi,K.
  TITLE     Secretory protein or membrane protein
  JOURNAL   Patent: EP 1514933-A 466 16-MAR-2005;
            Research Association for Biotechnology (JP)
FEATURES              Location/Qualifiers
     source           1. .726
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match            53.1%;  Score 685.2;  DB 6;  Length 726;
  Best Local Similarity  98.3%;  Pred. No. 1.2e-153;
  Matches  712;  Conservative    0;  Mismatches   10;  Indels    2;  Gaps    2;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||   | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCG-AAT 482
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAAAT 480

Qy        483 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 540

Qy        543 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 602
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 600

Qy        603 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGAT 662
                |||||||||||||||||||||||||||||||||||||||||||||||||||| | ||||||
Db        601 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCAAAAGGGAT 660

Qy        663 GTGTGTCAT-CGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAG 721
                ||||||||| ||||||| |||||| |||||||||||||||||||||||||||||| ||||||
Db        661 GTGTGTCATCCGGTGTAAGCACAANCGGTGGCACTTTATAAAGCCCACTAACANGTCCAG 720

Qy        722 AGAG 725
                ||||
Db        721 AGAG 724


RESULT 10
AX136544
LOCUS       AX136544                726 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 466 from Patent EP1067182.
ACCESSION   AX136544
VERSION     AX136544.1  GI:14272948
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
            Hayashi,K.
  TITLE     Secretory protein or membrane protein
  JOURNAL   Patent: EP 1067182-A 466 10-JAN-2001;
            Helix Research Institute (JP)
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           53.1%;  Score 685.2;  DB 6;  Length 726;
  Best Local Similarity   98.3%;  Pred. No. 1.2e-153;
  Matches  712;  Conservative    0;  Mismatches   10;  Indels    2;  Gaps    2;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCG-AAT 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAAAT 480

Qy        483 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 540

Qy        543 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 600

Qy        603 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGAT 662
              |||||||||||||||||||||||||||||||||||||||||||||||||||| | ||||||
Db        601 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCAAAAGGGAT 660

Qy        663 GTGTGTCAT-CGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAG 721
              ||||||||| ||||||| |||||| |||||||||||||||||||||||||||||| ||||||
Db        661 GTGTGTCATCCGGTGTAAGCACAANCGGTGGCACTTTATAAAGCCCACTAACANGTCCAG 720

Qy        722 AGAG 725
              ||||
Db        721 AGAG 724


RESULT 11
CQ727837
LOCUS       CQ727837                680 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 13771 from Patent WO02068579.
ACCESSION   CQ727837
VERSION     CQ727837.1  GI:42294978
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 13771 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           52.7%;  Score 680;  DB 6;  Length 680;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-152;
  Matches  680;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         93 ATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGC 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGC 60

Qy        153 GCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCCAGA 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCCAGA 120

Qy        213 ACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATATTGCA 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATATTGCA 180

Qy        273 TACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTG 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTG 240

Qy        333 CTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAG 392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 CTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAG 300

Qy        393 GTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATC 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATC 360

Qy        453 GTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGAT 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGAT 420

Qy        513 AACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTG 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTG 480

Qy        573 AGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTG 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTG 540

Qy        633 CATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGG 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGG 600

Qy        693 CACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTC 752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTC 660

Qy        753 ACGGTCCTTTCTGTTGGCAG 772
              ||||||||||||||||||||
Db        661 ACGGTCCTTTCTGTTGGCAG 680
```

```
RESULT 12
BC031198
LOCUS       BC031198                2581 bp    mRNA    linear   ROD 30-JUN-2004
DEFINITION  Mus musculus expressed sequence AA536743, mRNA (cDNA clone
            MGC:37691 IMAGE:5059693), complete cds.
ACCESSION   BC031198
VERSION     BC031198.1  GI:21410498
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2581)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2581)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 58 Row: n Column: 7
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 22122334.
FEATURES             Location/Qualifiers
     source          1. .2581
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="MGC:37691 IMAGE:5059693"
                     /tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
                     old mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam2"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .2581
                     /gene="AA536743"
                     /db_xref="GeneID:100532"
                     /db_xref="MGI:2140767"
     CDS             344. .952
                     /gene="AA536743"
                     /codon_start=1
                     /product="expressed sequence AA536743"
                     /protein_id="AAH31198.1"
                     /db_xref="GI:21410499"
                     /db_xref="GeneID:100532"
                     /db_xref="MGI:2140767"
                     /translation="MGLLGVLICHLLKKKGYRCTTEAEQEVEEEKVEKIELNDSINEN
                     SDTVGQIVQYIMKNEANADILKAMVADNSVGDIESPVTPSTPGSPPVSPGPLSPGATP
                     GKHVCGHHLHTVGGVVERDVCQRCRHKRWHFIKPTNKTKEGRPRRQGEVTVLSVGRFR
                     VTKVEHKSNQKERRSLMSVSGIESVNGDVPATPVKRERSDTE"
ORIGIN

  Query Match            48.2%;  Score 622.2;  DB 9;  Length 2581;
  Best Local Similarity  80.8%;  Pred. No. 1.9e-138;
  Matches  739;  Conservative    0;  Mismatches  173;  Indels    3;  Gaps    1;

Qy         14 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGG 73
              |||| | || |  |||| || | | || |||   | |||||||| ||| || ||| ||||
Db         55 CCCGGCAGCCCTCGAGCTGCACCGTCTACCCCGACCCGGGACGCCGCGGCCGGGGCGAGG 114

Qy         74 AGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGC 133
              |||||  || |||||||||||||||| | ||| |||||||||||||| || || || ||
Db        115 AGGGCAGAGGGACGCGGCGATGGCTCTGTGGGGACTCCCGGGGTCCGCTGTGCTTGCTGC 174

Qy        134 TGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAG 193
              | | || |||||||||||||| |||||||||||||||||||  | |||||     | ||
Db        175 TTCAGTGTTCGTGGGAGGCGCTGTGAGTTCGCCGCTGGTGGCCGCCGACAACACTGGTAG 234

Qy        194 CCGCACATTGCACTCCAGAACAGAGACGACCC---CGTCGCCCAGCAACGATACTGGGAA 250
              || ||| || |||||| || ||||||| || |   |||| |||| |||| |  |||||||
Db        235 CCACACGTTACACTCCCGAGCAGAGACCACTCCATCGTCACCCACCAACAACCCTGGGAA 294

Qy        251 TGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTT 310
              |||||||| ||||| || || |||| ||| |||||||||||||  |||||||||| |
Db        295 TGGACACCCGGAATACATCGCGTACGTGCTCGTCCCTGTGTTCTTCGTCATGGGTCTCCT 354
```

```
Qy        311 TGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGA 370
              ||| |||||||| || ||| | |||||||||||||||| || |||||||||| || ||
Db        355 TGGTGTCCTCATCTGTCACTTACTTAAGAAGAAAGGCTACCGGTGTACAACAGAGGCGGA 414

Qy        371 GCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAA 430
              || ||  |||||||||||||||| |||||||||| |||||||||||| | ||||| ||
Db        415 ACAGGAAGTCGAAGAGGAAAAGGTGGAAAAGATAGAGTTGAATGACAGTATAAATGAGAA 474

Qy        431 CAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGT 490
              ||||||||| || || || || ||||| ||||| |||||||||||||||||| ||| |
Db        475 CAGTGACACGGTCGGCCAGATTGTCCAGTACATTATGAAAAATGAAGCGAATGCAGATAT 534

Qy        491 CTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAG 550
              ||| || || |||||||| ||||||||| ||   |||  ||||||||| || || || ||
Db        535 CTTGAAAGCTATGGTAGCTGATAACAGCGTGGGCGATATTGAAAGCCCTGTCACTCCTAG 594

Qy        551 CACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAA 610
              ||| |||||||||||| || ||||||||||||||  |||||||||||||  || |||||||
Db        595 CACGCCAGGGAGCCCACCTGTGAGTCCTGGGCCCCTGTCACCAGGGGCCACCCCAGGGAA 654

Qy        611 GCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCA 670
              ||| |||||||| |||||||||| || ||||| ||||| || ||  ||||||||||| ||
Db        655 GCATGTCTGTGGTCATCATCTGCACACAGTGGGAGGTGTCGTTGAACGGGATGTGTGCCA 714

Qy        671 TCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAG 730
              || || |||||||||||| |||||||| ||||| ||||| |||||| ||| |||| | |
Db        715 GCGATGCAGGCACAAGCGATGGCACTTCATAAAACCCACCAACAAGACCAAAGAGGGACG 774

Qy        731 ACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGT 790
              |||||||| |||||||||||||| ||||| ||||| ||||| || ||||| || |||||
Db        775 GCCACGGCGACAAGGCGAGGTCACCGTCCTCTCTGTGGGCAGGTTCAGAGTCACGAAAGT 834

Qy        791 GGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAAC 850
              |||||||||||||| |||||||| || |||||  |||||||||  |||||   ||| |
Db        835 GGAGCACAAGTCAAATCAGAAGGAGCGCAGAAGTTTGATGTCTGTCAGTGGCATTGAGAG 894

Qy        851 CGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAGAGTAGCA 910
              |||||  ||||| ||||| || || |||||||||| || || |||| ||| |||||||
Db        895 TGTCAACGGGGATGTGCCTGCGACGCCTGTGAAGAGGGAGCGGAGTGACACGGAGTAGCT 954

Qy        911 GGTGAGCCGTGGTTT 925
              |  |   |  | |||
Db        955 GCAGCTGCTGGCTTT 969
```

```
RESULT 13
BC066137
LOCUS       BC066137               2481 bp    mRNA    linear   ROD 08-MAR-2005
DEFINITION  Mus musculus expressed sequence AA536743, mRNA (cDNA clone
            MGC:76429 IMAGE:30012041), complete cds.
ACCESSION   BC066137
VERSION     BC066137.1  GI:42490934
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2481)
  AUTHORS   Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
            Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
            B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
            Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin
            GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
            Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P,
            Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy
            SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
            Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,
            Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
            Helton E, Ketteman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
            Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW,
            Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,
            Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U,
            Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2481)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Minoru Ko
            cDNA Library Preparation: Yulan Piao and Minoru Ko (National
            Institute on Aging, NIH: http://lgsun.grc.nia.nih.gov/cDNA/)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www.shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 144 Row: a Column: 18.
FEATURES             Location/Qualifiers
     source          1. .2481
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="MGC:76429 IMAGE:30012041"
                     /tissue_type="Embryo, whole, mouse, 7.5-dpc"
                     /clone_lib="NIA Mouse 7.5-dpc Whole Embryo cDNA Library
                     (Long)"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1"
     gene            1. .2481
                     /gene="AA536743"
                     /db_xref="GeneID:100532"
                     /db_xref="MGI:2140767"
     CDS             354. .962
                     /gene="AA536743"
                     /codon_start=1
                     /product="expressed sequence AA536743"
                     /protein_id="AAH66137.1"
                     /db_xref="GI:42490935"
                     /db_xref="GeneID:100532"
                     /db_xref="MGI:2140767"
                     /translation="MGLLGVLICHLLKKKGYRCTTEAEQEVEEEKVEKIELNDSINEN
                     SDTVGQIVQYIMKNEANADILKAMVADNSVGDIESPVTPSTPGSPPVSPGPLSPGATP
                     GKHVCGHHLHTVGGVVERDVCQRCRHKRWHFIKPTNKTKEGRPRRQGEVTVLSVGRFR
                     VTKVEHKSNQKERRSLMSVSGIESVNGDVPATPVKRERSDTE"
ORIGIN

  Query Match           48.1%;  Score 620.6;  DB 9;  Length 2481;
  Best Local Similarity   80.7%;  Pred. No. 4.7e-138;
  Matches  738;  Conservative    0;  Mismatches  174;  Indels    3;  Gaps    1;

Qy         14 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGG 73
              |||| | || |  |||| || | | || |||    | ||||||||| ||| || ||| ||||
Db         65 CCCGGCAGCCCTCGAGCTGCACCGTCTACCCCGACCCGGGACGCCGCGGCCGGGGCGAGG 124

Qy         74 AGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGC 133
              |||||  || |||||||||||||||| | ||| |||||||||||||| || || || ||
Db        125 AGGGCAGAGGGACGCGGCGATGGCTCTGTGGGGACTCCCGGGGTCCGCTGTGCTTGCTGC 184

Qy        134 TGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAG 193
```

```
Db        185 TTCAGTGTTCGTGGGAGGCGCTGTGAGTTCGCCGCTGGTGGCCGCCGACAACACTGGTAG 244

Qy        194 CCGCACATTGCACTCCAGAACAGAGACGACCC---CGTCGCCCAGCAACGATACTGGGAA 250

Db        245 CCACACGTTACACTCCCGAGCAGAGACCACTCCATCGTCACCCACCAACAACCCTGGGAA 304

Qy        251 TGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTT 310

Db        305 TGGACACCCGGAATACATCGCGTACGTGCTCGTCCCTGTGTTCTTCGTCATGGGTCTCCT 364

Qy        311 TGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGA 370

Db        365 TGGTGTCCTCATCTGTCACTTACTTAAGAAGAAAGGCTACCGGTGTACAACGGAGGCGGA 424

Qy        371 GCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAA 430

Db        425 ACAGGAAGTCGAAGAGGAAAAGGTGGAAAAGATAGAGTTGAATGACAGTATAAATGAGAA 484

Qy        431 CAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGT 490

Db        485 CAGTGACACGGTCGGCCAGATTGTCCAGTACATTATGAAAAATGAAGCGAATGCAGATAT 544

Qy        491 CTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAG 550

Db        545 CTTGAAAGCTATGGTAGCTGATAACAGCGTGGGCGATATTGAAAGCCCTGTCACTCCTAG 604

Qy        551 CACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAA 610

Db        605 CACGCCAGGGAGCCCACCTGTGAGTCCTGGGCCCCTGTCACCAGGGGCCACCCCAGGGAA 664

Qy        611 GCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCA 670

Db        665 GCATGTCTGTGGTCATCATCTGCACACAGTGGGAGGTGTCGTTGAACGGGATGTGTGCCA 724

Qy        671 TCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAG 730

Db        725 GCGATGCAGGCACAAGCGATGGCACTTCATAAAACCCACCAACAAGACCAAAGAGGGACG 784

Qy        731 ACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGT 790

Db        785 GCCACGGCGACAAGGCGAGGTCACCGTCCTCTCTGTGGGCAGGTTCAGAGTCACGAAAGT 844

Qy        791 GGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAAC 850

Db        845 GGAGCACAAGTCAAATCAGAAGGAGCGCAGAAGTTTGATGTCTGTCAGTGGCATTGAGAG 904

Qy        851 CGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAGAGTAGCA 910

Db        905 TGTCAACGGGGATGTGCCTGCGACGCCTGTGAAGAGGGAGCGGAGTGACACGGAGTAGCT 964

Qy        911 GGTGAGCCGTGGTTT 925

Db        965 GCAGCTGCTGGCTTT 979


RESULT 14
BC066160
LOCUS       BC066160               2481 bp    mRNA    linear   ROD 08-MAR-2005
DEFINITION  Mus musculus expressed sequence AA536743, mRNA (cDNA clone
            MGC:76510 IMAGE:30012671), complete cds.
ACCESSION   BC066160
VERSION     BC066160.1  GI:42490882
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2481)
  AUTHORS   Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
            Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
            B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
            Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin
            GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
            Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P,
            Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy
            SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
            Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,
            Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
            Helton E, Ketteman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
            Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW,
            Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,
            Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U,
            Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   2  (bases 1 to 2481)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Minoru Ko
            cDNA Library Preparation: Yulan Piao and Minoru Ko (National
            Institute on Aging, NIH: http://lgsun.grc.nia.nih.gov/cDNA/)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 144 Row: a Column: 19
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 22122334.
FEATURES             Location/Qualifiers
     source          1. .2481
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="MGC:76510 IMAGE:30012671"
                     /tissue_type="Embryo, whole, mouse, 7.5-dpc"
                     /clone_lib="NIA Mouse 7.5-dpc Whole Embryo cDNA Library
                     (Long)"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1"
     gene            1. .2481
                     /gene="AA536743"
                     /db_xref="GeneID:100532"
                     /db_xref="MGI:2140767"
     CDS             354. .962
                     /gene="AA536743"
                     /codon_start=1
                     /product="expressed sequence AA536743"
                     /protein_id="AAH66160.1"
                     /db_xref="GI:42490883"
                     /db_xref="GeneID:100532"
                     /db_xref="MGI:2140767"
                     /translation="MGLLGVLICHLLKKKGYRCTTEAEQEVEEEKVEKIELNDSINEN
                     SDTVGQIVQYIMKNEANADILKAMVADNSVGDIESPVTPSTPGSPPVSPGPLSPGATP
                     GKHVCGHHLHTVGGVVERDVCQRCRHKRWHFIKPTNKTKEGRPRRQGEVTVLSVGRFR
                     VTKVEHKSNQKERRSLMSVSGIESVNGDVPATPVKRERSDTE"
ORIGIN
```

```
  Query Match           48.1%;  Score 620.6;  DB 9;  Length 2481;
  Best Local Similarity  80.7%;  Pred. No. 4.7e-138;
  Matches  738;  Conservative    0;  Mismatches  174;  Indels    3;  Gaps    1;

Qy         14 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGG 73
              |||| | || |  |||| || | | || |||    | ||||||| ||| || ||| ||||
Db         65 CCCGGCAGCCCTCGAGCTGCACCGTCTACCCCGACCCGGGACGCCGCGGCCGGGGCGAGG 124

Qy         74 AGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGC 133
              |||||  || ||||||||||||||||| | ||| ||||||||||||||| || || || ||
Db        125 AGGGCAGAGGGACGCGGCGATGGCTCTGTGGGGACTCCCGGGGTCCGCTGTGCTTGCTGC 184

Qy        134 TGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAG 193
              | | || ||||||||||||||| ||||||||||||||||||||  | |||||    | ||
Db        185 TTCAGTGTTCGTGGGAGGCGCTGTGAGTTCGCCGCTGGTGGCCGCCGACAACACTGGTAG 244

Qy        194 CCGCACATTGCACTCCAGAACAGAGACGACCC---CGTCGCCCAGCAACGATACTGGGAA 250
              || ||| || |||||| || ||||||| || |   |||| |||| |||| |  |||||||
Db        245 CCACACGTTACACTCCCGAGCAGAGACCACTCCATCGTCACCCACCAACAACCCTGGGAA 304

Qy        251 TGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTT 310
              ||||||||| ||||| || || |||| ||| ||||||||||||||  |||||||||| |
Db        305 TGGACACCCGGAATACATCGCGTACGTGCTCGTCCCTGTGTTCTTCGTCATGGGTCTCCT 364

Qy        311 TGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGA 370
              ||| |||||||| || ||| | |||||||||||||||| || |||||||| || || ||
Db        365 TGGTGTCCTCATCTGTCACTTACTTAAGAAGAAAGGCTACCGGTGTACAACGGAGGCGGA 424

Qy        371 GCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAA 430
              || ||  |||||||||||||||| |||||||||| |||||||||||  | ||||| ||
Db        425 ACAGGAAGTCGAAGAGGAAAAGGTGGAAAAGATAGAGTTGAATGACAGTATAAATGAGAA 484

Qy        431 CAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGT 490
              ||||||||| || || || || ||||| ||||| ||||||||||||||||||| ||| |
Db        485 CAGTGACACGGTCGGCCAGATTGTCCAGTACATTATGAAAAATGAAGCGAATGCAGATAT 544

Qy        491 CTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAG 550
              ||| || || |||||||| |||||||| ||   |||  |||||||| || || || ||
Db        545 CTTGAAAGCTATGGTAGCTGATAACAGCGTGGGCGATATTGAAAGCCCTGTCACTCCTAG 604

Qy        551 CACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAA 610
              ||| |||||||||||  || ||||||||||||||   ||||||||||||   || ||||||||
Db        605 CACGCCAGGGAGCCCACCTGTGAGTCCTGGGCCCCTGTCACCAGGGGCCACCCCAGGGAA 664

Qy        611 GCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCA 670
              ||| |||||||| |||||||||||| || ||||| ||||| || ||  |||||||||| ||
Db        665 GCATGTCTGTGGTCATCATCTGCACACAGTGGGAGGTGTCGTTGAACGGGATGTGTGCCA 724

Qy        671 TCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAG 730
              || || |||||||||||  |||||||| ||||| ||||| |||||| ||| |||| |  |
Db        725 GCGATGCAGGCACAAGCGATGGCACTTCATAAAACCCACCAACAAGACCAAAGAGGGACG 784

Qy        731 ACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGT 790
              ||||||||  |||||||||||||||| ||||| ||||| ||||| || ||||| || |||||
Db        785 GCCACGGCGACAAGGCGAGGTCACCGTCCTCTCTGTGGGCAGGTTCAGAGTCACGAAAGT 844

Qy        791 GGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAAC 850
              |||||||||||||| |||||||| || |||||  |||||||| |||||   ||| |
Db        845 GGAGCACAAGTCAAATCAGAAGGAGCGCAGAAGTTTGATGTCTGTCAGTGGCATTGAGAG 904

Qy        851 CGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAGAGTAGCA 910
              ||||| ||||| ||||| || || |||||||||| || || |||| ||| |||||||
Db        905 TGTCAACGGGGATGTGCCTGCGACGCCTGTGAAGAGGGAGCGGAGTGACACGGAGTAGCT 964

Qy        911 GGTGAGCCGTGGTTT 925
              |  |   |  | |||
Db        965 GCAGCTGCTGGCTTT 979


RESULT 15
AC027607/c
LOCUS       AC027607               161070 bp    DNA     linear   PRI 03-JUL-2002
DEFINITION  Homo sapiens BAC clone RP11-36B15 from 4, complete sequence.

ACCESSION   AC027607
VERSION     AC027607.7  GI:21686941
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 161070)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 161070)
  AUTHORS   Cordes,M., Cotton,M., Bradshaw-Cordum,H. and Boyer,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-36B15
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 161070)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 161070)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Jul 3, 2002 this sequence version replaced gi:15321552.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0036B15
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-562N2, 2000 bp overlap; the
            clone sequenced to the right is RP11-9D8, 2000 bp overlap.  Actual
            start of this clone is at base position 1 of RP11-36B15; actual end
            is at base position 161070 of RP11-36B15.
```

A transposon was identified in the vector portion of the clone.

There is an unresolved tandem repeat from base 71075 to 71587.

```
FEATURES             Location/Qualifiers
     source          1. .161070
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /map="4"
                     /clone="RP11-36B15"
                     /clone_lib="RPCI-11"
     repeat_region   160. .301
                     /rpt_family="MaLR"
     repeat_region   577. .942
                     /rpt_family="MaLR"
     repeat_region   943. .1287
                     /rpt_family="MaLR"
     repeat_region   1278. .1458
                     /rpt_family="MaLR"
     repeat_region   1471. .1508
                     /rpt_family="(CTATA)n"
     repeat_region   1517. .1622
                     /rpt_family="MER1_type"
     repeat_region   2137. .2274
                     /rpt_family="MER1_type"
     repeat_region   2440. .2540
                     /rpt_family="MER1_type"
     repeat_region   2907. .3188
                     /rpt_family="Alu"
     repeat_region   3189. .3251
                     /rpt_family="(TA)n"
     repeat_region   6539. .6848
                     /rpt_family="Alu"
     repeat_region   7980. .9503
                     /rpt_family="L1"
     repeat_region   9624. .9960
                     /rpt_family="MaLR"
     repeat_region   10113. .10602
                     /rpt_family="MaLR"
     repeat_region   10847. .10868
                     /rpt_family="(GGGAA)n"
     repeat_region   12489. .13469
                     /rpt_family="MER2_type"
     repeat_region   13470. .13773
                     /rpt_family="Alu"
     repeat_region   13774. .13969
                     /rpt_family="MER2_type"
     repeat_region   15751. .15870
                     /rpt_family="MIR"
     repeat_region   16222. .16530
                     /rpt_family="Alu"
     repeat_region   16627. .16942
                     /rpt_family="Alu"
     repeat_region   17033. .17180
                     /rpt_family="MIR"
     repeat_region   17419. .17542
                     /rpt_family="AcHobo"
     repeat_region   17543. .17847
                     /rpt_family="Alu"
     repeat_region   17848. .18547
                     /rpt_family="AcHobo"
     repeat_region   18875. .19044
                     /rpt_family="L1"
     repeat_region   19045. .19093
                     /rpt_family="AT_rich"
     repeat_region   20459. .20819
                     /rpt_family="MaLR"
     repeat_region   20860. .20930
                     /rpt_family="L2"
     repeat_region   21174. .21330
                     /rpt_family="Alu"
     repeat_region   21663. .21764
                     /rpt_family="MIR"
     repeat_region   21791. .21979
                     /rpt_family="MIR"
     repeat_region   22247. .22480
                     /rpt_family="MaLR"
     repeat_region   22568. .22730
                     /rpt_family="MER1_type"
     repeat_region   23743. .23902
                     /rpt_family="MIR"
     repeat_region   23912. .23991
                     /rpt_family="(CACCAT)n"
     repeat_region   24000. .24024
                     /rpt_family="(CA)n"
     repeat_region   24905. .25306
                     /rpt_family="ERVL"
     repeat_region   25307. .25599
                     /rpt_family="Alu"
     repeat_region   25600. .25697
                     /rpt_family="ERVL"
     repeat_region   25805. .25928
                     /rpt_family="L1"
     repeat_region   27584. .27687
                     /rpt_family="MIR"
     repeat_region   27703. .28168
                     /rpt_family="ERVL"
     repeat_region   28471. .28675
                     /rpt_family="MER1_type"
     repeat_region   28810. .28902
                     /rpt_family="L2"
     repeat_region   29062. .29161
                     /rpt_family="L1"
     repeat_region   29261. .29491
                     /rpt_family="L1"
     repeat_region   29492. .29780
                     /rpt_family="Alu"
     repeat_region   29781. .30593
                     /rpt_family="L1"
     repeat_region   30777. .31087
                     /rpt_family="Alu"
     repeat_region   31143. .31520
                     /rpt_family="L1"
     repeat_region   31973. .32016
                     /rpt_family="(ATG)n"
     repeat_region   32121. .32601
                     /rpt_family="ERV1"
     repeat_region   32602. .32991
                     /rpt_family="ERV1"
     repeat_region   32963. .33107
                     /rpt_family="ERV1"
     repeat_region   33088. .33711
                     /rpt_family="ERV1"
     repeat_region   33697. .34700
```

```
  Query Match             40.2%;  Score 519.6;  DB 8;  Length 161070;
  Best Local Similarity   98.3%;  Pred. No. 1.4e-113;
  Matches  525;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        758 CCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACG 817
              || |||      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117470 CCATTCCCCCCTCAGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACG 117411

Qy        818 GAGAAGCCTGATGTCTGTTAGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACC 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117410 GAGAAGCCTGATGTCTGTTAGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACC 117351

Qy        878 TGTGAAGAGAGAACGCAGTGGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGG 937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117350 TGTGAAGAGAGAACGCAGTGGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGG 117291

Qy        938 GGGCAGAGTGGTGCAGGGTGAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCA 997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117290 GGGCAGAGTGGTGCAGGGTGAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCA 117231
```

```
Qy        998 TAGGAATGGTATTTGACAGGGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGG 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117230 TAGGAATGGTATTTGACAGGGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGG 117171

Qy       1058 GACTGAACATGATTACTTGTCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGC 1117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117170 GACTGAACATGATTACTTGTCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGC 117111

Qy       1118 CTCCAGGGGCTTGGCCTGTGTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGC 1177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117110 CTCCAGGGGCTTGGCCTGTGTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGC 117051

Qy       1178 ATGGTGGGGATTGTGGCAAACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACC 1237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117050 ATGGTGGGGATTGTGGCAAACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACC 116991

Qy       1238 ACGAGGCAGAAAGCACCTAGTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||||||||||||||||||||||| | |||
Db     116990 ACGAGGCAGAAAGCACCTAGTGTTGCTGCATCTTCTTACGCAAAAAAAGACAAA 116937
```

Search completed: February 23, 2006, 17:05:11
Job time : 6923 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 23, 2006, 14:34:42 ; Search time 786 Seconds
(without alignments)
10946.715 Million cell updates/sec

Title: US-10-620-562-1
Perfect score: 1291
Sequence: 1 cccacgcgtccggcccgccg.....cttacgcaaaaaaaaaaaaa 1291

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1291	100.0	1291	6	AAD33366	Aad33366 Human TR2
	2	1196.6	92.7	1603	2	AAV84456	Aav84456 Human sec
	3	1196.6	92.7	1603	4	ABA83239	Aba83239 Human sec
	4	1196.6	92.7	1603	9	ACH04740	Ach04740 Novel hum
	5	1196.6	92.7	1603	9	ACD44550	Acd44550 Human cDN
	6	900	69.7	1320	5	AAF93833	Aaf93833 Human cDN
	7	900	69.7	1320	14	ADY63174	Ady63174 Human clo
	8	891	69.0	1218	10	ADI21815	Adi21815 Novel hum
	9	827.8	64.1	861	4	AAH98631	Aah98631 Human EST
	10	827.8	64.1	861	10	ADI21359	Adi21359 Novel hum
	11	685.2	53.1	726	5	AAF94032	Aaf94032 Primer sp
	12	685.2	53.1	726	14	ADY63459	Ady63459 Human clo
	13	172.6	13.4	348	13	ADU02019	Adu02019 Novel hum
	14	115.4	8.9	694	6	ABQ19006	Abq19006 Oligonucl
c	15	115.4	8.9	694	6	ABQ19007	Abq19007 Oligonucl
c	16	104.6	8.1	694	6	ABQ19008	Abq19008 Oligonucl
	17	104.6	8.1	694	6	ABQ19009	Abq19009 Oligonucl
	18	73	5.7	488	9	ACH13424	Ach13424 Human adu
	19	72.6	5.6	582	6	AAD33367	Aad33367 Human TR2
	20	72.6	5.6	827	4	AAK92116	Aak92116 Human cDN
	21	72.6	5.6	827	4	AAK93378	Aak93378 Human cDN
	22	72.6	5.6	827	12	ADL29805	Adl29805 5' end of
	23	72.6	5.6	827	12	ADL28543	Adl28543 5' end of
	24	72.6	5.6	1254	9	ACC59914	Acc59914 Human REM
	25	72.6	5.6	1409	4	AAK94427	Aak94427 Human ful
	26	72.6	5.6	1409	12	ADL31171	Adl31171 Full leng
	27	72.6	5.6	1793	8	ABX70777	Abx70777 Novel hum
c	28	72.6	5.6	2886	4	ABK43492	Abk43492 DNA encod
c	29	72.6	5.6	2886	12	ADI53879	Adi53879 cDNA enco
	30	58.6	4.5	1290	2	AAZ23121	Aaz23121 Human TAN
	31	58.6	4.5	1293	6	ABS53807	Abs53807 Human TR4
	32	58.6	4.5	1293	14	ADV43505	Adv43505 Human psy
	33	58.6	4.5	1860	12	ADH22385	Adh22385 Human cDN
	34	58.6	4.5	2570	2	AAZ23120	Aaz23120 Human TAN
c	35	58.6	4.5	2576	4	AAF45119	Aaf45119 Human sec
	36	58.6	4.5	2638	4	AAD10353	Aad10353 Human tum
	37	58.6	4.5	2703	3	AAD00061	Aad00061 Human tum
	38	58.6	4.5	2703	4	AAS15031	Aas15031 DNA encod
	39	58.6	4.5	2732	8	ADA83982	Ada83982 Human TNF
	40	58.6	4.5	2732	13	ACN39280	Acn39280 Tumour-as
	41	58.6	4.5	3374	14	ADZ13507	Adz13507 Human can
	42	58.6	4.5	3402	13	ABD33404	Abd33404 Human can
	43	58.6	4.5	3407	14	ADZ13509	Adz13509 Human can
	44	58.6	4.5	3434	14	ADZ13515	Adz13515 Human can
c	45	58.6	4.5	3531	4	AAD08454	Aad08454 Human sec

ALIGNMENTS

RESULT 1
AAD33366
ID AAD33366 standard; cDNA; 1291 BP.
XX
AC AAD33366;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human TR21 receptor cDNA.
XX
KW Human; tumour necrosis factor receptor; TNF; TR21; TR22; immune disorder;
KW autoimmune haemolytic anaemia; rheumatoid arthritis; Addison's disease;
KW allergy; cancer; ulcerative colitis; cardiovascular disorder; epilepsy;
KW myocardial ischaemia; wound healing; neurological disease; infection;
KW cerebral anoxia; gene therapy; immunosuppressive; antiproliferative;
KW cytostatic; cardiant; vasotropic; neuroprotective; antibacterial;
KW virucide; fungicide; nootropic; ophthalmological; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 93. .908
FT /*tag= a
FT /product= "Human TR21 protein"
FT sig_peptide 93. .179
FT /*tag= b
FT mat_peptide 180. .905
FT /*tag= c
FT /product= "Human mature TR21 protein"
XX
PN WO200207762-A1.
XX
PD 31-JAN-2002.
XX
PF 23-JUL-2001; 2001WO-US023124.
XX
PR 24-JUL-2000; 2000US-0220116P.
PR 27-JUL-2000; 2000US-0221143P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Zeng Z, Ruben SM, Rosen CA;

XX
DR WPI; 2002-171948/22.
DR P-PSDB; AAE20962.
XX
PT Two novel proteins, TR21 and TR22, which are members of the tumor
PT necrosis factor receptor, useful for the diagnosis and treatment of
PT immune disorders, cancer, cardiovascular disorders.
XX
PS Claim 1; Fig 1; 248pp; English.
XX
CC The present invention relates to novel human tumour necrosis factor (TNF)
CC receptors, TR21 and TR22 and polynucleotides encoding them. Sequences of
CC the invention are useful in the diagnosis, treatment and prevention of
CC cancers (e.g., cancers of the adrenal gland, bone, urogenital or bone
CC marrow, in particular breast and ovarian cancer), immune disorders (e.g.,
CC autoimmune haemolytic anaemia, rheumatoid arthritis, allergies, Addison's
CC disease, ulcerative colitis), cardiovascular disorders (e.g., myocardial
CC ischaemias), wound healing, neurological diseases (e.g., cerebral anoxia,
CC epilepsy) and infectious diseases such as viral, bacterial, fungal and
CC parasitic infections. They are also useful in gene therapy. The present
CC sequence is a cDNA encoding human TR21 protein
XX
SQ Sequence 1291 BP; 326 A; 302 C; 405 G; 258 T; 0 U; 0 Other;

Query Match 100.0%; Score 1291; DB 6; Length 1291;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAA 360

Qy        361 CAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTG 420

Qy        421 TGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGA 480

Qy        481 ATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCG 540

Qy        541 TGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGA 600

Qy        601 CGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGG 660

Qy        661 ATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCA 720

Qy        721 GAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAG 780

Qy        781 TTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTG 840

Qy        841 GGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCA 900

Qy        901 CAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGTGAGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGTGAGG 960

Qy        961 AGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAGGGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAGGGAA 1020

Qy       1021 GTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTGTCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTGTCTG 1080

Qy       1081 CCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGTGTGA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGTGTGA 1140

Qy       1141 TAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAAACTA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAAACTA 1200

Qy       1201 GAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTAGTGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTAGTGT 1260

Qy       1261 TGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              ||||||||||||||||||||||||||||||
Db       1261 TGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
```

RESULT 2
AAV84456
ID AAV84456 standard; DNA; 1603 BP.
XX
AC AAV84456;
XX
DT 01-MAR-1999 (first entry)
XX
DE Human secreted protein gene 46 clone HCFMV39.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9854963-A2.
XX
PD 10-DEC-1998.
XX
PF 04-JUN-1998; 98WO-US011422.
XX
PR 06-JUN-1997; 97US-0048875P.
PR 06-JUN-1997; 97US-0048876P.

PR 06-JUN-1997; 97US-0048877P.
PR 06-JUN-1997; 97US-0048878P.
PR 06-JUN-1997; 97US-0048880P.
PR 06-JUN-1997; 97US-0048881P.
PR 06-JUN-1997; 97US-0048882P.
PR 06-JUN-1997; 97US-0048883P.
PR 06-JUN-1997; 97US-0048884P.
PR 06-JUN-1997; 97US-0048885P.
PR 06-JUN-1997; 97US-0048892P.
PR 06-JUN-1997; 97US-0048893P.
PR 06-JUN-1997; 97US-0048894P.
PR 06-JUN-1997; 97US-0048895P.
PR 06-JUN-1997; 97US-0048896P.
PR 06-JUN-1997; 97US-0048897P.
PR 06-JUN-1997; 97US-0048898P.
PR 06-JUN-1997; 97US-0048899P.
PR 06-JUN-1997; 97US-0048900P.
PR 06-JUN-1997; 97US-0048901P.
PR 06-JUN-1997; 97US-0048915P.
PR 06-JUN-1997; 97US-0048916P.
PR 06-JUN-1997; 97US-0048917P.
PR 06-JUN-1997; 97US-0048949P.
PR 06-JUN-1997; 97US-0048962P.
PR 06-JUN-1997; 97US-0048963P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048970P.
PR 06-JUN-1997; 97US-0048971P.
PR 06-JUN-1997; 97US-0048972P.
PR 06-JUN-1997; 97US-0048974P.
PR 06-JUN-1997; 97US-0049019P.
PR 06-JUN-1997; 97US-0049020P.
PR 06-JUN-1997; 97US-0049373P.
PR 06-JUN-1997; 97US-0049374P.
PR 06-JUN-1997; 97US-0049375P.
PR 05-SEP-1997; 97US-0057584P.
PR 05-SEP-1997; 97US-0057627P.
PR 05-SEP-1997; 97US-0057628P.
PR 05-SEP-1997; 97US-0057629P.
PR 05-SEP-1997; 97US-0057634P.
PR 05-SEP-1997; 97US-0057635P.
PR 05-SEP-1997; 97US-0057642P.
PR 05-SEP-1997; 97US-0057643P.
PR 05-SEP-1997; 97US-0057644P.
PR 05-SEP-1997; 97US-0057645P.
PR 05-SEP-1997; 97US-0057646P.
PR 05-SEP-1997; 97US-0057647P.
PR 05-SEP-1997; 97US-0057648P.
PR 05-SEP-1997; 97US-0057649P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057651P.
PR 05-SEP-1997; 97US-0057654P.
PR 05-SEP-1997; 97US-0057661P.
PR 05-SEP-1997; 97US-0057662P.
PR 05-SEP-1997; 97US-0057666P.
PR 05-SEP-1997; 97US-0057667P.
PR 05-SEP-1997; 97US-0057668P.
PR 05-SEP-1997; 97US-0057760P.
PR 05-SEP-1997; 97US-0057761P.
PR 05-SEP-1997; 97US-0057762P.
PR 05-SEP-1997; 97US-0057763P.
PR 05-SEP-1997; 97US-0057764P.
PR 05-SEP-1997; 97US-0057765P.
PR 05-SEP-1997; 97US-0057769P.
PR 05-SEP-1997; 97US-0057770P.
PR 05-SEP-1997; 97US-0057771P.
PR 05-SEP-1997; 97US-0057774P.
PR 05-SEP-1997; 97US-0057775P.
PR 05-SEP-1997; 97US-0057776P.
PR 05-SEP-1997; 97US-0057777P.
PR 05-SEP-1997; 97US-0057778P.
PR 18-DEC-1997; 97US-0070923P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;
PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;
PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;
PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;
PI Carter KC;
XX
DR WPI; 1999-059865/05.
DR P-PSDB; AAW88579.
XX
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
PS Claim 4; Page 315-316; 772pp; English.
XX
CC The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
CC encoding human secreted proteins (AAW88534 to AAW88756). The secreted
CC protein gene sequences are deposited with the ATCC under deposit numbers
CC ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners. The
CC present sequence represents a gene encoding a human secreted protein (see
CC descriptor line for gene number and clone identification)
XX
SQ Sequence 1603 BP; 416 A; 356 C; 449 G; 371 T; 0 U; 11 Other;

Query Match 92.7%; Score 1196.6; DB 2; Length 1603;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1268; Conservative 5; Mismatches 13; Indels 9; Gaps 6;

```
Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||| |  :     | ||||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
```

```
                |||||||||||||||||||||||||| | ||||||||||||||:|||:||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
                |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
                |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
                |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
                ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
                |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
                |||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296


RESULT 3
ABA83239
ID   ABA83239 standard; cDNA; 1603 BP.
XX
AC   ABA83239;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Human secreted protein gene 46 SEQ ID NO:56.
XX
KW   Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
KW   dermatological; immunosuppressive; antiinflammatory; immunostimulant;
KW   cytostatic; cardiant; vascular; anti-angiogenic; ophthalmological;
KW   neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnerary;
KW   antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW   multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW   human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW   Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis;
KW   Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder;
KW   corneal graft neovascularisation; diabetic retinopathy; regeneration;
KW   neurological disorder; Huntington's chorea; Alzheimer's disease;
KW   Parkinson's disease; infectious disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162891-A2.
XX
PD   30-AUG-2001.
XX
PF   21-FEB-2001; 2001WO-US005614.
XX
PR   24-FEB-2000; 2000US-0184836P.
PR   29-MAR-2000; 2000US-0193170P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J,  Ebner R,  Lafleur DW,  Moore PA,  Olsen HS,  Rosen CA;
PI   Ruben SM,  Soppet DR,  Young PE,  Shi Y,  Florence KA,  Wei Y;
PI   Florence C,  Hu J,  Li Y,  Kyaw H,  Fischer CL,  Ferrie AM,  Fan P;
PI   Feng P,  Endress GA,  Dillon PJ,  Carter KC,  Brewer LA,  Yu G,  Zeng Z;
PI   Greene JM;
XX
DR   WPI; 2001-625724/72.
DR   P-PSDB; ABB50346.
XX
PT   Nucleic acids encoding 207 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease
PT   and diabetic retinopathy.
XX
PS   Claim 1; Page 944-945; 1533pp; English.
XX
CC   ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
CC   proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
CC   activities based on the tissues and cells the genes are expressed in.
CC   Example of these activities include: immunomodulatory; antisclerotic;
CC   dermatological; immunosuppressive; antiinflammatory; immunostimulant;
CC   anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;
CC   neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;
CC   antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used
CC   in gene therapy and vaccine production. (I) and (II) can be used in the
CC   prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC   sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC   (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC   Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
CC   Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC   disorders (e.g. corneal graft neovascularisation and diabetic
CC   retinopathy), neurological disorders (e.g. Huntington's chorea,
CC   Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC   for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
CC   ABA83193 and ABB50300 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1603 BP; 416 A; 356 C; 449 G; 371 T; 0 U; 11 Other;

  Query Match           92.7%;  Score 1196.6;  DB 4;  Length 1603;
  Best Local Similarity   97.9%;  Pred. No. 0;
  Matches 1268;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    6;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64
```

```
Qy        61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy       121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy       181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy       241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy       301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
             |||||||||||||||||||||||||  |  :     | ||||||||||||||||||||||
Db       304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy       358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
             |||||||||||||||||||||||||||| | ||||||||||||||:|||:||||||||||
Db       364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy       417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy       477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
             ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db       484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy       537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy       597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy       657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy       717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy       777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy       837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy       897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy       957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy      1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy      1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db      1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy      1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy      1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy      1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
             |||||||||||||||||||||||||||| | ||||
Db      1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296
```

RESULT 4
ACH04740
ID ACH04740 standard; cDNA; 1603 BP.
XX
AC ACH04740;
XX
DT 02-OCT-2003 (first entry)
XX
DE Novel human secreted protein #46 cDNA.
XX
KW Human; ss; gene therapy; autoimmune disorder; multiple sclerosis; cancer;
KW systemic lupus erythematosus; haematopoietic cell disorder; allergy;
KW agammaglobulinaemia; ataxia telangiectasia; blood coagulation disorder;
KW afibrinogenaemia; thrombocytopenia; graft-versus-host disease; arthritis;
KW inflammatory condition; ischaemia-reperfusion injury; infectious disease;
KW hyperproliferative disorder; purpura; viral infection; regeneration;
KW bacterial infection; ulcer; Alzheimer's disease; gene.
XX
OS Homo sapiens.
XX
PN US2003065160-A1.
XX
PD 03-APR-2003.
XX
PF 07-DEC-2001; 2001US-00004860.
XX
PR 06-JUN-1997; 97US-0048875P.
PR 06-JUN-1997; 97US-0048876P.
PR 06-JUN-1997; 97US-0048877P.
PR 06-JUN-1997; 97US-0048878P.
PR 06-JUN-1997; 97US-0048880P.
PR 06-JUN-1997; 97US-0048881P.
PR 06-JUN-1997; 97US-0048882P.
PR 06-JUN-1997; 97US-0048883P.
PR 06-JUN-1997; 97US-0048884P.
PR 06-JUN-1997; 97US-0048885P.
PR 06-JUN-1997; 97US-0048892P.
PR 06-JUN-1997; 97US-0048893P.
PR 06-JUN-1997; 97US-0048894P.
PR 06-JUN-1997; 97US-0048895P.
PR 06-JUN-1997; 97US-0048896P.
PR 06-JUN-1997; 97US-0048897P.
PR 06-JUN-1997; 97US-0048898P.
PR 06-JUN-1997; 97US-0048899P.
PR 06-JUN-1997; 97US-0048900P.
PR 06-JUN-1997; 97US-0048901P.
PR 06-JUN-1997; 97US-0048915P.
PR 06-JUN-1997; 97US-0048916P.
PR 06-JUN-1997; 97US-0048917P.
PR 06-JUN-1997; 97US-0048949P.
PR 06-JUN-1997; 97US-0048962P.
PR 06-JUN-1997; 97US-0048963P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048970P.
PR 06-JUN-1997; 97US-0048971P.
PR 06-JUN-1997; 97US-0048972P.
PR 06-JUN-1997; 97US-0048974P.
PR 06-JUN-1997; 97US-0049019P.
PR 06-JUN-1997; 97US-0049020P.
PR 06-JUN-1997; 97US-0049373P.
PR 06-JUN-1997; 97US-0049374P.

```
PR   06-JUN-1997;   97US-0049375P.
PR   05-SEP-1997;   97US-0057584P.
PR   05-SEP-1997;   97US-0057627P.
PR   05-SEP-1997;   97US-0057628P.
PR   05-SEP-1997;   97US-0057629P.
PR   05-SEP-1997;   97US-0057634P.
PR   05-SEP-1997;   97US-0057635P.
PR   05-SEP-1997;   97US-0057642P.
PR   05-SEP-1997;   97US-0057643P.
PR   05-SEP-1997;   97US-0057644P.
PR   05-SEP-1997;   97US-0057645P.
PR   05-SEP-1997;   97US-0057646P.
PR   05-SEP-1997;   97US-0057647P.
PR   05-SEP-1997;   97US-0057648P.
PR   05-SEP-1997;   97US-0057649P.
PR   05-SEP-1997;   97US-0057650P.
PR   05-SEP-1997;   97US-0057651P.
PR   05-SEP-1997;   97US-0057654P.
PR   05-SEP-1997;   97US-0057661P.
PR   05-SEP-1997;   97US-0057662P.
PR   05-SEP-1997;   97US-0057666P.
PR   05-SEP-1997;   97US-0057667P.
PR   05-SEP-1997;   97US-0057668P.
PR   05-SEP-1997;   97US-0057760P.
PR   05-SEP-1997;   97US-0057761P.
PR   05-SEP-1997;   97US-0057762P.
PR   05-SEP-1997;   97US-0057763P.
PR   05-SEP-1997;   97US-0057764P.
PR   05-SEP-1997;   97US-0057765P.
PR   05-SEP-1997;   97US-0057769P.
PR   05-SEP-1997;   97US-0057770P.
PR   05-SEP-1997;   97US-0057771P.
PR   05-SEP-1997;   97US-0057774P.
PR   05-SEP-1997;   97US-0057775P.
PR   05-SEP-1997;   97US-0057776P.
PR   05-SEP-1997;   97US-0057777P.
PR   05-SEP-1997;   97US-0057778P.
PR   18-DEC-1997;   97US-0070923P.
PR   04-JUN-1998;   98WO-US011422.
PR   15-JUL-1998;   98US-0092921P.
PR   30-JUL-1998;   98US-0094657P.
PR   04-DEC-1998;   98US-00205258.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Young P,  Greene JM,  Ferrie AM,  Ruben SM,  Rosen CA,  Hu J;
PI   Olsen HS,  Ebner R,  Brewer LA,  Moore PA,  Shi Y,  Florence C;
PI   Florence K,  Lafleur DW,  Ni J,  Fan P,  Wei Y,  Fischer CL,  Soppet DR;
PI   Li Y,  Zeng Z,  Kyaw H,  Yu G,  Peng P,  Dillon PJ,  Endress GA;
PI   Carter KC;
XX
DR   WPI; 2003-540804/51.
DR   P-PSDB; ABO44603.
XX
PT   New isolated protein, useful for preparing a composition for diagnosing
PT   or treating cancer, inflammatory, immune or infectious diseases.
XX
PS   Example 1; SEQ ID NO 56; 172pp; English.
XX
CC   The invention relates to an isolated HEMAE80 protein. The protein is
CC   useful for preparing a composition for diagnosing or treating autoimmune
CC   disorders e.g. multiple sclerosis and systemic lupus erythematosus;
CC   haematopoietic cell disorders e.g. agammaglobulinaemia and ataxia
CC   telangiectasia; blood coagulation disorders e.g. afibrinogenaemia and
CC   thrombocytopenia; allergy; graft-versus-host disease; inflammatory
CC   conditions e.g. ischaemia-reperfusion injury and arthritis;
CC   hyperproliferative disorders e.g. cancer and purpura; infectious disease
CC   e.g. viral infection and bacterial infection. The polynucleotide or
CC   protein can be used to regenerate damaged tissue e.g. ulcers and
CC   Alzheimer's disease. The present sequence represents a novel human
CC   secreted protein cDNA. Note: The sequence data for this patent did not
CC   form part of the printed specification but was obtained in electronic
CC   format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030065160
XX
SQ   Sequence 1603 BP; 416 A; 356 C; 449 G; 371 T; 0 U; 11 Other;

  Query Match            92.7%;  Score 1196.6;  DB 9;  Length 1603;
  Best Local Similarity  97.9%;  Pred. No. 0;
  Matches 1268;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    6;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||||||  |  :      | ||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||||| | |||||||||||||:|||:||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
                |||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296
```

RESULT 5
ACD44550
ID ACD44550 standard; cDNA; 1603 BP.
XX
AC ACD44550;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human cDNA from novel secreted protein gene 46.
XX
KW Human; ss; gene; secreted protein; precerebellin-like protein;
KW neurodegenerative disorder; behavioural disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia; mania;
KW dementia; paranoia; psychosis; autism; immune disorder; infection;
KW inflammation; allergy; liver disorder; hepatoblastoma; jaundice;
KW hepatitis; immunological disorder; AIDS; leukaemia; rheumatoid arthritis;
KW sepsis; acne; psoriasis; cancer.
XX
OS Homo sapiens.
XX
PN US6525174-B1.
XX
PD 25-FEB-2003.
XX
PF 04-DEC-1998; 98US-00205258.
XX
PR 06-JUN-1997; 97US-0048875P.
PR 06-JUN-1997; 97US-0048876P.
PR 06-JUN-1997; 97US-0048877P.
PR 06-JUN-1997; 97US-0048878P.
PR 06-JUN-1997; 97US-0048880P.
PR 06-JUN-1997; 97US-0048881P.
PR 06-JUN-1997; 97US-0048882P.
PR 06-JUN-1997; 97US-0048883P.
PR 06-JUN-1997; 97US-0048884P.
PR 06-JUN-1997; 97US-0048885P.
PR 06-JUN-1997; 97US-0048892P.
PR 06-JUN-1997; 97US-0048893P.
PR 06-JUN-1997; 97US-0048894P.
PR 06-JUN-1997; 97US-0048895P.
PR 06-JUN-1997; 97US-0048896P.
PR 06-JUN-1997; 97US-0048897P.
PR 06-JUN-1997; 97US-0048898P.
PR 06-JUN-1997; 97US-0048899P.
PR 06-JUN-1997; 97US-0048900P.
PR 06-JUN-1997; 97US-0048901P.
PR 06-JUN-1997; 97US-0048915P.
PR 06-JUN-1997; 97US-0048916P.
PR 06-JUN-1997; 97US-0048917P.
PR 06-JUN-1997; 97US-0048949P.
PR 06-JUN-1997; 97US-0048962P.
PR 06-JUN-1997; 97US-0048963P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048970P.
PR 06-JUN-1997; 97US-0048971P.
PR 06-JUN-1997; 97US-0048972P.
PR 06-JUN-1997; 97US-0048974P.
PR 06-JUN-1997; 97US-0049019P.
PR 06-JUN-1997; 97US-0049020P.
PR 06-JUN-1997; 97US-0049373P.
PR 06-JUN-1997; 97US-0049374P.
PR 06-JUN-1997; 97US-0049375P.
PR 05-SEP-1997; 97US-0057584P.
PR 05-SEP-1997; 97US-0057627P.
PR 05-SEP-1997; 97US-0057628P.
PR 05-SEP-1997; 97US-0057629P.
PR 05-SEP-1997; 97US-0057634P.
PR 05-SEP-1997; 97US-0057635P.
PR 05-SEP-1997; 97US-0057642P.
PR 05-SEP-1997; 97US-0057643P.
PR 05-SEP-1997; 97US-0057644P.
PR 05-SEP-1997; 97US-0057645P.
PR 05-SEP-1997; 97US-0057646P.
PR 05-SEP-1997; 97US-0057647P.
PR 05-SEP-1997; 97US-0057648P.
PR 05-SEP-1997; 97US-0057649P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057651P.
PR 05-SEP-1997; 97US-0057654P.
PR 05-SEP-1997; 97US-0057661P.
PR 05-SEP-1997; 97US-0057662P.
PR 05-SEP-1997; 97US-0057666P.
PR 05-SEP-1997; 97US-0057667P.
PR 05-SEP-1997; 97US-0057668P.
PR 05-SEP-1997; 97US-0057760P.
PR 05-SEP-1997; 97US-0057761P.
PR 05-SEP-1997; 97US-0057762P.
PR 05-SEP-1997; 97US-0057763P.
PR 05-SEP-1997; 97US-0057764P.
PR 05-SEP-1997; 97US-0057765P.
PR 05-SEP-1997; 97US-0057769P.
PR 05-SEP-1997; 97US-0057770P.
PR 05-SEP-1997; 97US-0057771P.
PR 05-SEP-1997; 97US-0057774P.
PR 05-SEP-1997; 97US-0057775P.
PR 05-SEP-1997; 97US-0057776P.
PR 05-SEP-1997; 97US-0057777P.
PR 05-SEP-1997; 97US-0057778P.
PR 18-DEC-1997; 97US-0070923P.
PR 04-JUN-1998; 98WO-US011422.
PR 15-JUL-1998; 98US-0092921P.
PR 30-JUL-1998; 98US-0094657P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;
PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;
PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;
PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;
PI Carter KC;
XX
DR WPI; 2003-511926/48.
DR P-PSDB; ABO26083.
XX
PT New precerebellin-like protein, useful for diagnosing or treating
PT neurodegenerative and behavioral disorders, immune disorders, liver
PT disorders, and cancer.
XX
PS Example 1; SEQ ID NO 56; 156pp; English.
XX
CC The invention relates to an isolated protein comprising amino acid
CC residues 33-205 or 1-205 of a a novel human secreted protein appearing as

```
CC   ABO26252. The protein is encoded by one of 238 disclosed cDNA sequences
CC   encoding 238 secreted proteins. ABO26252 is a precerebellin-like protein.
CC   Also included are a composition comprising the protein and a carrier and
CC   an isolated protein produced by expressing the protein cited above by a
CC   cell, and recovering the protein. The proteins are useful for diagnosing
CC   or treating neurodegenerative and behavioural disorders (e.g. Alzheimer's
CC   disease, Parkinson's disease, Huntington's disease, schizophrenia, mania,
CC   dementia, paranoia, psychoses or autism), immune disorders (e.g.
CC   infection, inflammation, allergy), liver disorders (e.g. hepatoblastoma,
CC   jaundice, hepatitis), immunological disorders (e.g. AIDS, leukaemia,
CC   rheumatoid arthritis, sepsis, acne, psoriasis) and cancer. The present
CC   sequence is one of the 238 disclosed cDNAs encoding a novel secreted
CC   protein. Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from USPTO at:- seqdata.uspto.gov/sequence.html?DocID=6525174B1
XX
SQ   Sequence 1603 BP; 416 A; 356 C; 449 G; 371 T; 0 U; 11 Other;

  Query Match           92.7%;  Score 1196.6;  DB 9;  Length 1603;
  Best Local Similarity  97.9%;  Pred. No. 0;
  Matches 1268;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    6;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||||  |  :      | ||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||| | |||||||||||||:|||:||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||  ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296


RESULT 6
AAF93833
ID   AAF93833 standard; cDNA; 1320 BP.
XX
AC   AAF93833;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Human cDNA encoding a membrane or secretory protein clone PSEC0162.
XX
KW   Human; secretory protein; membrane protein; vaccine; gene therapy;
KW   rheumatoid arthritis; diabetes; ss.
XX
OS   Homo sapiens.
XX
PN   EP1067182-A2.
XX
PD   10-JAN-2001.
XX
PF   07-JUL-2000; 2000EP-00114090.
XX
PR   08-JUL-1999;   99JP-00194179.
PR   11-JAN-2000; 2000JP-00118775.
PR   02-MAY-2000; 2000JP-00183766.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Kawai Y,  Sugiyama T,  Hayashi K;
XX
DR   WPI; 2001-093989/11.
DR   P-PSDB; AAB88406.
XX
PT   Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT   gene therapy or as candidate target molecules in drug development.
XX
```

```
PS   Claim 1; SEQ ID NO 179; 609pp + Sequence Listing; English.
XX
CC   This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC   which encode human secretory or membrane proteins represented by AAB88317
CC   - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and
CC   AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the
CC   invention. The invention also includes methods for the production of
CC   antibodies directed against the proteins, and cDNA sequences, which can
CC   be used in vaccines. The polynucleotide sequences can be used in gene
CC   therapy. The polynucleotide sequences and the proteins they encode may be
CC   used in the prevention, treatment and diagnosis of diseases associated
CC   with inappropriate secretory protein/membrane protein expression. The
CC   nucleic acids and complementary sequences may also be used as DNA probes
CC   in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect
CC   and quantitate the presence of similar nucleic acid sequences in samples.
CC   They may also be used to study the expression and function of secretory
CC   proteins/membrane polypeptides and their role in metabolism. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   against them and in assays to identify modulators (agonists and
CC   antagonists) of expression and activity. The antibodies and antagonists
CC   may also be used as therapeutic agents to down regulate expression and
CC   activity. The antibodies may also be used as diagnostic agents for
CC   detecting the presence of the polypeptides in samples (e.g. by enzyme
CC   linked immunosorbant assay (ELISA). Examples of diseases which may be
CC   treated include rheumatoid arthritis and diabetes
XX
SQ   Sequence 1320 BP; 332 A; 359 C; 370 G; 259 T; 0 U; 0 Other;

  Query Match             69.7%;  Score 900;  DB 5;  Length 1320;
  Best Local Similarity   99.4%;  Pred. No. 4.5e-251;
  Matches  903;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              ||  ||   | ||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 480

Qy        484 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 540

Qy        544 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 600

Qy        604 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 660

Qy        664 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 720

Qy        724 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 780

Qy        784 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 840

Qy        844 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 900

Qy        904 AGTAGCAG 911
              ||||||||
Db        901 AGTAGCAG 908


RESULT 7
ADY63174
ID   ADY63174 standard; cDNA; 1320 BP.
XX
AC   ADY63174;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Human clone PSEC0162 cDNA, SEQ ID 179.
XX
KW   Gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   EP1514933-A1.
XX
PD   16-MAR-2005.
XX
PF   07-JUL-2000; 2004EP-00027228.
XX
PR   08-JUL-1999;   99JP-00194179.
PR   11-JAN-2000; 2000JP-00118775.
PR   02-MAY-2000; 2000JP-00183766.
PR   07-JUL-2000; 2000EP-00114090.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Kawai Y,  Sugiyama T,  Hayashi K;
XX
DR   WPI; 2005-203865/22.
DR   P-PSDB; ADY63175.
XX
PT   Novel isolated polynucleotide encoding human secretary proteins or
PT   membrane proteins, useful for examination and diagnosis of abnormality of
PT   human secretary proteins.
XX
PS   Disclosure; SEQ ID NO 179; 1240pp; English.
XX
CC   The present invention relates to novel human secretory proteins or
CC   membrane proteins, and their coding sequences. The present sequence is
CC   one such coding sequence. The coding sequences of the invention are
CC   useful for examination and diagnosis of abnormality of the human
CC   secretary proteins and in gene therapy methods. The coding sequences and
CC   proteins are useful as candidates for medicines or as target molecules
CC   for developing medicines. Antibodies against the proteins of the
CC   invention are useful for treating diseases that are associated with the
CC   proteins. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained from sequence information
CC   supplied by the European Patent Office.
```

```
XX
SQ   Sequence 1320 BP; 332 A; 359 C; 370 G; 259 T; 0 U; 0 Other;

  Query Match            69.7%;  Score 900;  DB 14;  Length 1320;
  Best Local Similarity  99.4%;  Pred. No. 4.5e-251;
  Matches  903;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              ||  ||   | ||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATG 480

Qy        484 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGA 540

Qy        544 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGC 600

Qy        604 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATG 660

Qy        664 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAG 720

Qy        724 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTA 780

Qy        784 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGG 840

Qy        844 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAG 900

Qy        904 AGTAGCAG 911
              ||||||||
Db        901 AGTAGCAG 908


RESULT 8
```

```
ADI21815
ID   ADI21815 standard; cDNA; 1218 BP.
XX
AC   ADI21815;
XX
DT   15-APR-2004  (first entry)
XX
DE   Novel human protein cDNA #74.
XX
KW   forensic; nutritional source; damaged tissue; diseased tissue;
KW   myeloid cell disorder; lymphoid cell disorder;
KW   bone cartilage tissue growth; tendon tissue growth;
KW   ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW   tissue repair; tissue replacement; burn; incision; ulcer; cancer; human;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO2003025148-A2.
XX
PD   27-MAR-2003.
XX
PF   19-SEP-2002; 2002WO-US029964.
XX
PR   19-SEP-2001; 2001US-0323739P.
PR   13-SEP-2002; 2002US-00323739.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Asundi V,  Goodrich RW,  Ren F,  Zhang J,  Zhao QA,  Wang J;
PI   Ghosh M,  Xue AJ,  Wehrman T,  Weng G,  Zhou P,  Drmanac RT,  Wang D;
PI   Haley-Vicente D;
XX
DR   WPI; 2003-354603/33.
DR   P-PSDB; ADI21099.
XX
PT   New polynucleotides and secreted proteins, useful for treating myeloid or
PT   lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT   tissue growth or regeneration, in wound healing, and in tissue repair and
PT   replacement.
XX
PS   Claim 1; SEQ ID NO 74; 156pp; English.
XX
CC   The invention relates to an isolated polynucleotide encoding a
CC   polypeptide with biological activity. The polynucleotides and
CC   polypeptides are useful in diagnostics, forensics, gene mapping,
CC   identification of mutations responsible for genetic disorders and other
CC   traits, to assess biodiversity, as nutritional sources or supplements.
CC   The polynucleotides may also be used as molecular weight markers,
CC   chromosome markers or map related gene positions, or as an antigen to
CC   raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC   useful for raising antibodies, as markers for tissues in which the
CC   corresponding polypeptide is expressed, for re-engineering damaged or
CC   diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC   bone cartilage, tendon, ligament and/or nerve tissue growth or
CC   regeneration, in wound healing, in tissue repair and replacement, in
CC   healing of burns, incisions and ulcers, and in treating cancer. The
CC   present sequence represents cDNA encoding a novel human protein.
XX
SQ   Sequence 1218 BP; 318 A; 302 C; 368 G; 230 T; 0 U; 0 Other;

  Query Match            69.0%;  Score 891;  DB 10;  Length 1218;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-248;
  Matches  891;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 GCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGA 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGA 60

Qy         81 AGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTC 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTC 120
```

```
Qy        141 TTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACA 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACA 180

Qy        201 TTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCA 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCA 240

Qy        261 GAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTC 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTC 300

Qy        321 ATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATC 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATC 360

Qy        381 GAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACT 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACT 420

Qy        441 GTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCG 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCG 480

Qy        501 ATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGG 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGG 540

Qy        561 AGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGT 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGT 600

Qy        621 GGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGG 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGG 660

Qy        681 CACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGC 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGC 720

Qy        741 CAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAG 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAG 780

Qy        801 TCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAACCGTCAATGGG 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAACCGTCAATGGG 840

Qy        861 GAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAGAGTAGCAG 911
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAGAGTAGCAG 891


RESULT 9
AAH98631
ID   AAH98631 standard; cDNA; 861 BP.
XX
AC   AAH98631;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Human EST-derived coding sequence SEQ ID NO: 488.
XX
KW   Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW   tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW   diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
KW   gene therapy; nutrition; ss.
XX
OS   Homo sapiens.
XX
PN   WO200154477-A2.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-US002687.
XX
PR   25-JAN-2000; 2000US-00491404.
PR   17-JUL-2000; 2000US-00617746.
PR   03-AUG-2000; 2000US-00631451.
PR   15-SEP-2000; 2000US-00663870.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Qian XB,  Wang Z,  Chen R,  Asundi V;
PI   Cao Y,  Drmanac RA,  Zhang J,  Werhman T;
XX
DR   WPI; 2001-476164/51.
DR   P-PSDB; AAM23972.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use.
XX
PS   Claim 1; Page 522; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a cDNA
CC   of the invention
XX
SQ   Sequence 861 BP; 216 A; 214 C; 265 G; 166 T; 0 U; 0 Other;

  Query Match           64.1%;  Score 827.8;  DB 4;  Length 861;
  Best Local Similarity  98.5%;  Pred. No. 3.9e-230;
  Matches  846;  Conservative    0;  Mismatches   12;  Indels    1;  Gaps    1;

Qy         30 CGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCG 89
              || ||||  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CGACTCTTTCGTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCG 61

Qy         90 GCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGA 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGA 121

Qy        150 GGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCC 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCC 181

Qy        210 AGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATATT 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATATT 241

Qy        270 GCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCAC 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCAC 301

Qy        330 CTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAA 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAA 361

Qy        390 AAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAA 449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAA 421

Qy        450 ATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCA 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCA 481

Qy        510 GATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCA-GGGAGCCCGCC 568
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        482 GATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGGAGCCCGCC 541

Qy        569 AGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATCA 628
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 AGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATCA 601

Qy        629 TCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAAGCG 688
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 TCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAAGCG 661

Qy        689 GTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGCGA 748
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGCGA 721

Qy        749 GGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAACCA 808
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAACCA 781

Qy        809 GAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAACCGTCAATGGGGAGGTGCC 868
                |||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||
Db        782 CAAGGAACGGAGAAGCCTGATGTCTGTTAATGGGGCTGAAACCGTCCATGGGGAGGTGCC 841

Qy        869 GGCAACACCTGTGAAGAGA 887
                ||||||| ||    |||||
Db        842 GGCAACAACTTGTGAGAGA 860
```

RESULT 10
ADI21359
ID ADI21359 standard; cDNA; 861 BP.
XX
AC ADI21359;
XX
DT 15-APR-2004 (first entry)
XX
DE Novel human expressed sequence tag, EST #58.
XX
KW forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human;
KW ss; gene; EST; expressed sequence tag.
XX
OS Homo sapiens.
XX
PN WO2003025148-A2.
XX
PD 27-MAR-2003.
XX
PF 19-SEP-2002; 2002WO-US029964.
XX
PR 19-SEP-2001; 2001US-0323739P.
PR 13-SEP-2002; 2002US-00323739.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI Haley-Vicente D;
XX
DR WPI; 2003-354603/33.
DR P-PSDB; ADI21579.
XX
PT New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.
XX
PS Example 2; SEQ ID NO 610; 156pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers,
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents a novel human expressed sequence tag, EST.
XX
SQ Sequence 861 BP; 216 A; 214 C; 265 G; 166 T; 0 U; 0 Other;

```
  Query Match             64.1%;  Score 827.8;  DB 10;  Length 861;
  Best Local Similarity   98.5%;  Pred. No. 3.9e-230;
  Matches  846;  Conservative    0;  Mismatches   12;  Indels    1;  Gaps    1;

Qy         30 CGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCG 89
              || ||||  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CGACTCTTTCGTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCG 61

Qy         90 GCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGA 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGA 121

Qy        150 GGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCC 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCC 181

Qy        210 AGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATATT 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATATT 241

Qy        270 GCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCAC 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCAC 301

Qy        330 CTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAA 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAA 361

Qy        390 AAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAA 449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAA 421

Qy        450 ATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCA 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCA 481

Qy        510 GATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCA-GGGAGCCCGCC 568
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        482 GATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGGAGCCCGCC 541

Qy        569 AGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATCA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 AGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATCA 601

Qy        629 TCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAAGCG 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 TCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAAGCG 661

Qy        689 GTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGCGA 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGCGA 721

Qy        749 GGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAACCA 808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAACCA 781

Qy        809 GAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAACCGTCAATGGGGAGGTGCC 868
              |||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||
Db        782 CAAGGAACGGAGAAGCCTGATGTCTGTTAATGGGGCTGAAACCGTCCATGGGGAGGTGCC 841
```

```
Qy        869 GGCAACACCTGTGAAGAGA 887
              ||||||| ||    |||||
Db        842 GGCAACAACTTGTGAGAGA 860


RESULT 11
AAF94032
ID   AAF94032 standard; DNA; 726 BP.
XX
AC   AAF94032;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Primer specific for DNA encoding secretory/membrane protein SEQ ID 466.
XX
KW   Human; secretory protein; membrane protein; vaccine; gene therapy;
KW   rheumatoid arthritis; diabetes; PCR primer; ss.
XX
OS   Synthetic.
XX
PN   EP1067182-A2.
XX
PD   10-JAN-2001.
XX
PF   07-JUL-2000; 2000EP-00114090.
XX
PR   08-JUL-1999;   99JP-00194179.
PR   11-JAN-2000; 2000JP-00118775.
PR   02-MAY-2000; 2000JP-00183766.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Kawai Y,  Sugiyama T,  Hayashi K;
XX
DR   WPI; 2001-093989/11.
XX
PT   Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT   gene therapy or as candidate target molecules in drug development.
XX
PS   Claim 4; SEQ ID NO 466; 609pp + Sequence Listing; English.
XX
CC   This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC   which encode human secretory or membrane proteins represented by AAB88317
CC   - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and
CC   AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the
CC   invention. The invention also includes methods for the production of
CC   antibodies directed against the proteins, and cDNA sequences, which can
CC   be used in vaccines. The polynucleotide sequences can be used in gene
CC   therapy. The polynucleotide sequences and the proteins they encode may be
CC   used in the prevention, treatment and diagnosis of diseases associated
CC   with inappropriate secretory protein/membrane protein expression. The
CC   nucleic acids and complementary sequences may also be used as DNA probes
CC   in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect
CC   and quantitate the presence of similar nucleic acid sequences in samples.
CC   They may also be used to study the expression and function of secretory
CC   proteins/membrane polypeptides and their role in metabolism. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   against them and in assays to identify modulators (agonists and
CC   antagonists) of expression and activity. The antibodies and antagonists
CC   may also be used as therapeutic agents to down regulate expression and
CC   activity. The antibodies may also be used as diagnostic agents for
CC   detecting the presence of the polypeptides in samples (e.g. by enzyme
CC   linked immunosorbant assay (ELISA). Examples of diseases which may be
CC   treated include rheumatoid arthritis and diabetes
XX
SQ   Sequence 726 BP; 174 A; 192 C; 219 G; 138 T; 0 U; 3 Other;

  Query Match             53.1%;  Score 685.2;  DB 5;  Length 726;
  Best Local Similarity   98.3%;  Pred. No. 1.2e-188;
  Matches  712;  Conservative    0;  Mismatches   10;  Indels    2;  Gaps    2;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCG-AAT 482
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAAAT 480

Qy        483 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 540

Qy        543 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 600

Qy        603 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGAT 662
              |||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||
Db        601 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCAAAAGGGAT 660

Qy        663 GTGTGTCAT-CGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAG 721
              |||||||||  ||||||| |||||| ||||||||||||||||||||||||| ||||||
Db        661 GTGTGTCATCCGGTGTAAGCACAANCGGTGGCACTTTATAAAGCCCACTAACANGTCCAG 720

Qy        722 AGAG 725
              ||||
Db        721 AGAG 724


RESULT 12
ADY63459
ID   ADY63459 standard; cDNA; 726 BP.
XX
AC   ADY63459;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Human clone PSEC0162 5'-end cDNA, SEQ ID 466.
XX
KW   Gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1514933-A1.
XX
PD   16-MAR-2005.
XX
PF   07-JUL-2000; 2004EP-00027228.
XX
PR   08-JUL-1999;   99JP-00194179.
PR   11-JAN-2000; 2000JP-00118775.
```

PR 02-MAY-2000; 2000JP-00183766.
PR 07-JUL-2000; 2000EP-00114090.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX
DR WPI; 2005-203865/22.
XX
PT Novel isolated polynucleotide encoding human secretary proteins or
PT membrane proteins, useful for examination and diagnosis of abnormality of
PT human secretary proteins.
XX
PS Example 1; SEQ ID NO 466; 1240pp; English.
XX
CC The present invention relates to novel human secretory proteins or
CC membrane proteins, and their coding sequences. The coding sequences of
CC the invention are useful for examination and diagnosis of abnormality of
CC the human secretary proteins and in gene therapy methods. The coding
CC sequences and proteins are useful as candidates for medicines or as
CC target molecules for developing medicines. Antibodies against the
CC proteins of the invention are useful for treating diseases that are
CC associated with the proteins. The present sequence is a 5'-end sequence
CC of one such coding sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained from sequence information supplied by the European Patent
CC Office.
XX
SQ Sequence 726 BP; 174 A; 192 C; 219 G; 138 T; 0 U; 3 Other;

Query Match 53.1%; Score 685.2; DB 14; Length 726;
Best Local Similarity 98.3%; Pred. No. 1.2e-188;
Matches 712; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

```
Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 60

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 120

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 180

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 243
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGAGCAGCCGCGCATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATA 240

Qy        244 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGG 300

Qy        304 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAG 360

Qy        364 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGA 420

Qy        424 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCG-AAT 482
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |
Db        421 ATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAAAT 480

Qy        483 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 540

Qy        543 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 600

Qy        603 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGAT 662
              |||||||||||||||||||||||||||||||||||||||||||||||||||| | ||||||
Db        601 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCAAAAGGGAT 660

Qy        663 GTGTGTCAT-CGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAG 721
              ||||||||| ||||||| |||||| |||||||||||||||||||||||||| ||||||
Db        661 GTGTGTCATCCGGTGTAAGCACAANCGGTGGCACTTTATAAAGCCCACTAACANGTCCAG 720

Qy        722 AGAG 725
              ||||
Db        721 AGAG 724
```

RESULT 13
ADU02019
ID ADU02019 standard; cDNA; 348 BP.
XX
AC ADU02019;
XX
DT 27-JAN-2005 (first entry)
XX
DE Novel human polynucleotide seqid 486.
XX
KW cytostatic; antipsoriatic; antiinflammatory; gene therapy; Nanodisc;
KW proliferative disorder; inflammatory disorder; immune disorder;
KW metabolic disorder; bone disorder; CNS disorder; cancer; psoriasis;
KW ulcerative colitis; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004093804-A2.
XX
PD 04-NOV-2004.
XX
PF 19-APR-2004; 2004WO-US012047.
XX
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PR 08-SEP-2003; 2003US-0505059P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Lee E, Hestir K, Chu K, Masuoka L, Williams LT;
XX
DR WPI; 2004-775861/76.
DR P-PSDB; ADU02751.
XX
PT New first nucleic acid molecule comprising a polynucleotide sequence
PT given in the specification, useful in preparing a composition for
PT diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
XX
PS Claim 1; SEQ ID NO 486; 291pp; English.
XX
CC The invention describes a new first nucleic acid molecule comprising a
CC polynucleotide sequence given in the specification. Also described are:
CC an animal injected with the nucleic acid molecule; a second nucleic acid
CC molecule comprising a second polynucleotide sequence that is at least
CC about 70, 80, 90 or 95% homologous to the first nucleic acid molecule or
CC that hybridises to the first polynucleotide sequence under high
CC stringency conditions; a vector comprising the nucleic acid molecule and
CC a promoter that drives the expression of the nucleic acid molecule; a
CC host cell transformed, transfected, transduced or infected with the

CC nucleic acid molecule; a nucleic acid composition comprising a carrier or
CC a buffer and one or more compositions comprising the nucleic acid
CC molecule, vector or host cell; a substantially purified polypeptide; an
CC animal injected with the polypeptide; a polypeptide composition
CC comprising the polypeptide molecule and a carrier or buffer; a cell
CC culture medium comprising the polypeptide or transfected cells
CC transfected with the polynucleotide; making a transformed, transfected,
CC transduced, or infected host cell; synthesising Nanodiscs simultaneously
CC and for synthesising a series of simultaneously-synthesised Nanodiscs
CC sequentially utilising a dynamic system; preparing a hydrophobic protein
CC for determination of crystal structure; immunising a non-human animal;
CC screening for modulators of hydrophobic protein activity; a diagnostic
CC kit; determining the presence of the nucleic acid molecule or its
CC complement; determining the presence of an antibody to the polypeptide in
CC a sample; an antibody specifically recognising, binding to or modulating
CC the biological activity of at least one polypeptide encoded by a nucleic
CC acid molecule or its biologically active fragment; an antibody
CC composition comprising the antibody and a carrier; a bacteriophage, where
CC the antibody is displayed on the bacteriophage; a bacterial cell
CC comprising the bacteriophage; a non-human animal injected with the
CC antibody composition; a host cell that secretes the antibody; making an
CC antibody; diagnosing a disease, disorder, syndrome, or condition
CC comprising cancer, or proliferative, inflammatory, immune, metabolic,
CC bone, CNS, genetic, bacterial and viral diseases, disorders, syndromes or
CC conditions in a patient; a modulator composition comprising a modulator
CC and a carrier; gene therapy; prophylactic or therapeutic treatment of a
CC subject; an isolated modified cell comprising at least one first
CC heterologous nucleic acid molecule, where the first heterologous nucleic
CC acid molecule comprises a first polynucleotide sequence that encodes a
CC first polypeptide; a non-human animal deficient in the polypeptide or
CC that over-expresses the polypeptide; isolated tissues derived from the
CC non-human animal; and one or more cells derived from the non-human
CC animal. The nucleic acid is useful in preparing a composition for
CC diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
CC This sequence encodes a novel human polypeptide of the invention.
XX
SQ Sequence 348 BP; 34 A; 121 C; 146 G; 47 T; 0 U; 0 Other;

Query Match 13.4%; Score 172.6; DB 13; Length 348;
Best Local Similarity 97.8%; Pred. No. 1.8e-39;
Matches 175; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 CCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 61
              |||| ||  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 CCACTCGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 101

Qy         62 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 161

Qy        122 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 220
```

RESULT 14
ABQ19006
ID ABQ19006 standard; DNA; 694 BP.
XX
AC ABQ19006;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 5597.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 694 BP; 74 A; 86 C; 302 G; 232 T; 0 U; 0 Other;

Query Match 8.9%; Score 115.4; DB 6; Length 694;
Best Local Similarity 81.2%; Pred. No. 1.2e-22;
Matches 134; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy         16 CGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAG 75
              || ||| | |||||||| | ||  ||  ||||||||||||||||||   ||||||||||
Db        246 CGTCGGTTTCGGAGCGGTTTTGTTTTTTCGAGCGCGGGACGCGGCGTTTTGGGGGAGGAG 305

Qy         76 GGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTG 135
              |||||||||||||||||||| | |||||| | |  ||||||  || |||  ||| || ||
Db        306 GGCGAAGCGACGCGGCGATGGTTTCGCGGGTATTTTCGGGGTTCGTCGTTTTAGTCGTTG 365

Qy        136 CTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||| ||||||||||||| |||||||||| || |||||| | |||
Db        366 TTGTTTTCGTGGGAGGCGTCGTGAGTTCGTCGTTGGTGGTTTCGG 410
```

RESULT 15
ABQ19007/c
ID ABQ19007 standard; DNA; 694 BP.
XX
AC ABQ19007;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 5598.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;

KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 694 BP; 232 A; 302 C; 86 G; 74 T; 0 U; 0 Other;

Query Match 8.9%; Score 115.4; DB 6; Length 694;
Best Local Similarity 81.2%; Pred. No. 1.2e-22;
Matches 134; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy         16 CGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAG 75
              || ||| | |||||||| | ||  ||  ||||||||||||||||||    ||||||||||
Db        449 CGTCGGTTTCGGAGCGGTTTTGTTTTTTCGAGCGCGGGACGCGGCGTTTTGGGGGAGGAG 390

Qy         76 GGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTG 135
              |||||||||||||||||||| | |||||| | |  |||||| || |||  ||| || ||
Db        389 GGCGAAGCGACGCGGCGATGGTTTCGCGGGTATTTTCGGGGTTCGTCGTTTTAGTCGTTG 330

Qy        136 CTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||| |||||||||||||| |||||||||| || |||||| | |||
Db        329 TTGTTTTCGTGGGAGGCGTCGTGAGTTCGTCGTTGGTGGTTTCGG 285
```

Search completed: February 23, 2006, 14:48:06
Job time : 791 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 23, 2006, 14:48:23 ; Search time 4935 Seconds
(without alignments)
12239.539 Million cell updates/sec

Title: US-10-620-562-1
Perfect score: 1291
Sequence: 1 cccacgcgtccggcccgccg.....cttacgcaaaaaaaaaaaaa 1291

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	855.2	66.2	898	5	BX463460	BX463460 BX463460
2	765.2	59.3	882	2	BE792154	BE792154 601582045
3	756.8	58.6	882	5	BX423892	BX423892 BX423892
4	726.8	56.3	1000	3	BM919548	BM919548 AGENCOURT
5	594.4	46.0	620	5	BQ807844	NISC_kk10
6	594.4	46.0	874	5	BU150311	BU150311 AGENCOURT
7	570	44.2	577	8	DR004573	DR004573 TC104761
8	554.2	42.9	815	8	DN948289	DN948289 AGENCOURT
9	538.2	41.7	684	3	BP438847	BP438847 BP438847
10	488.6	37.8	906	3	BI691144	BI691144 603314531
11	479.6	37.1	626	8	DN382824	LIB38534_
12	460.2	35.6	795	3	BI689796	BI689796 603316010
13	454.6	35.2	680	5	BY754410	BY754410 BY754410
14	454.2	35.2	757	3	BI658107	BI658107 603285885
15	449.4	34.8	672	5	BY754598	BY754598 BY754598
16	447.8	34.7	813	6	CA459450	CA459450 AGENCOURT
17	445.4	34.5	705	11	DQ043383	DQ043383 Homo sapi
18	439.4	34.0	710	5	BY735770	BY735770 BY735770
19	407.4	31.6	612	2	BB619564	BB619564 BB619564
20	405.6	31.4	705	11	DQ043384	DQ043384 Pan trogl
21	387.2	30.0	578	2	BE282422	BE282422 601102659
22	385.6	29.9	660	2	BB612565	BB612565 BB612565
23	384.4	29.8	661	2	BB624998	BB624998 BB624998
24	383.8	29.7	624	7	CF899514	CF899514 A0304B02-
25	380	29.4	873	6	CA458776	CA458776 AGENCOURT
26	368.4	28.5	645	2	BG975812	BG975812 602842535
27	367.6	28.5	652	5	BY727150	BY727150 BY727150
28	364.2	28.2	1013	2	BG862046	BG862046 602795059
c 29	359	27.8	441	1	AA782933	AA782933 ai62d12.s
30	336.4	26.1	722	7	CF950164	CF950164 UI-M-HJ0-
31	319	24.7	464	7	CK334279	CK334279 H3012F11-
32	318.4	24.7	476	1	AI593167	AI593167 vs86d09.y
33	315.2	24.4	466	5	BY256640	BY256640 BY256640
34	314.6	24.4	643	2	BI113590	BI113590 602897859
35	312.4	24.2	494	2	BF016184	BF016184 uy31h11.y
36	309.6	24.0	490	1	AI788006	AI788006 ul16b01.y
37	308.2	23.9	580	5	BY265553	BY265553 BY265553
38	306	23.7	779	3	BP156823	BP156823 BP156823
39	305	23.6	428	4	AK182372	AK182372 Mus muscu
40	288.6	22.4	973	5	BU234449	BU234449 603412374
41	284.2	22.0	475	2	BB857593	BB857593 BB857593
42	278.8	21.6	441	5	BY273890	BY273890 BY273890
43	271.4	21.0	389	4	AK202598	AK202598 Mus muscu
44	271.4	21.0	765	7	CO559467	CO559467 AGENCOURT
45	271	21.0	464	6	CA564115	CA564115 K0322C05-

ALIGNMENTS

RESULT 1
BX463460
LOCUS BX463460 898 bp mRNA linear EST 06-MAY-2004
DEFINITION BX463460 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone CS0DN002YB14 5-PRIME, mRNA sequence.
ACCESSION BX463460
VERSION BX463460.2 GI:47070031
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 898)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 22, 2003 this sequence version replaced gi:31033530.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 2573.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DN002DA07QP1&c=2573.r.
FEATURES Location/Qualifiers
source 1. .898
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN002YB14"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN

```
  Query Match            66.2%;  Score 855.2;  DB 5;  Length 898;
  Best Local Similarity  97.1%;  Pred. No. 8.8e-221;
  Matches  873;  Conservative   15;  Mismatches    9;  Indels    2;  Gaps    2;

Qy         14 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGG 73
                ||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||||
Db          1 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGGGNCGCCCTGGGGGAGG 60

Qy         74 AGGGCGAAGCGACGCGG-CGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCG 132
                ||||||||:    |||| ||||||||||||||||||||||||||||||||||||||||
Db         61 AGGGCGAASGACAGCGGNCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCG 120

Qy        133 CTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCA 192
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCA 180

Qy        193 GCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATG 252
                ||||||||:|||||||:|:||:|:||||||||||||||||||||||||||||||||||||
Db        181 GCCGCACAWTGCACTCMARAAMARAGACGACCCCGTCGCCCAGCAACGATACTGGGAATG 240

Qy        253 GACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTG 312
                ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAMACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTG 300

Qy        313 GCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGC 372
                ||||||||||||||||||||||||||||||||||||||||||:||:||||||||||||
Db        301 GCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTAMAAMAGAAGCAGAGC 360

Qy        373 AAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACA 432
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACA 420

Qy        433 GTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCT 492
                ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGAMACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCT 480

Qy        493 TAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCA 552
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCA 540

Qy        553 CACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGC 612
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGC 600

Qy        613 ACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATC 672
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        601 ACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTYATC 660

Qy        673 GGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGAC 732
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGAC 720

Qy        733 CACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGG 792
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGG 780

Qy        793 AGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAACCG 852
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGAAACCG 840

Qy        853 TCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGTGGCACAGAGTAGCAG 911
                |||||||||||||||||| |||||||||||| :|: :||:| ||||||||||||||
Db        841 TCAATGGGGAGGTGCCGGCAAAACCTGTGAAGAGA-RAMASAGKGCCACAGAGTAGCAG 898


RESULT 2
BE792154
LOCUS       BE792154                882 bp    mRNA    linear   EST 20-SEP-2000
DEFINITION  601582045F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3936491 5',
            mRNA sequence.
ACCESSION   BE792154
VERSION     BE792154.1  GI:10213352
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM780  row: j  column: 12
            High quality sequence stop: 825.
FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3936491"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_7"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN

  Query Match            59.3%;  Score 765.2;  DB 2;  Length 882;
  Best Local Similarity  99.5%;  Pred. No. 2.5e-196;
  Matches  778;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy         26 GGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGA 85
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGA 60

Qy         86 CGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGT 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGT 120

Qy        146 GGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCA 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCA 180

Qy        206 CTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATA 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATA 240

Qy        266 TATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTG 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTG 300

Qy        326 CCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGA 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGA 360

Qy        386 GGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGG 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGG 420

Qy        446 GCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGT 505
```

```
Db        421 GCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGT 480

Qy        506 AGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCC 565

Db        481 AGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCC 540

Qy        566 GCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCA 625

Db        541 GCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCA 600

Qy        626 TCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAA 685

Db        601 TCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAA 660

Qy        686 GCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGG 745

Db        661 GCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGG 720

Qy        746 CGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGA-GTTACAAAAGTGGAGCACAAGTCAA 804

Db        721 CGAGGTCACGGTCCTTTCTGTGGGCAGATTTAGAGGTTACAAAAGTGGAGCACAAGTTAA 780

Qy        805 AC 806

Db        781 CC 782
```

```
RESULT 3
BX423892
LOCUS       BX423892                882 bp    mRNA    linear   EST 01-MAY-2004
DEFINITION  BX423892 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN002YB14 5-PRIME, mRNA sequence.
ACCESSION   BX423892
VERSION     BX423892.2  GI:46932572
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30655585.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 2573.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS1AN001ZD08QP1&c=2573.r.
FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN002YB14"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN
```

```
  Query Match           58.6%;  Score 756.8;  DB 5;  Length 882;
  Best Local Similarity  96.7%;  Pred. No. 4.8e-194;
  Matches  803;  Conservative    9;  Mismatches   13;  Indels    5;  Gaps    4;

Qy         14 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGG 73

Db          1 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACG-GNCGCCCTGGGGGAGG 59

Qy         74 AGGGCGAAGCGACGCGG-CGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCG 132

Db         60 AGGGCGAASGANCGCGGNCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCG 119

Qy        133 CTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCA 192

Db        120 CTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCA 179

Qy        193 GCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATG 252

Db        180 GCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATG 239

Qy        253 GACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTG 312

Db        240 GACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTG 299

Qy        313 GCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGC 372

Db        300 GCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGC 359

Qy        373 AAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACA 432

Db        360 AAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACA 419

Qy        433 GTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCT 492

Db        420 GTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCT 479

Qy        493 TAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCA 552

Db        480 TAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCA 539

Qy        553 CACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGC 612

Db        540 CACCAGGGAGCCCGCCAGTGAGTTCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGC 599

Qy        613 ACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATC 672

Db        600 ACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTYATC 659

Qy        673 GGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGAC 732

Db        660 GGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGAC 719

Qy        733 CACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGG 792

Db        720 CACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTAC-AAARKKG 778

Qy        793 AGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTTAGTGGG 842

Db        779 RGCACAAGTCAACCAGAAAGGA--CGGGGAASCTGATKTYTGTTAGTGGG 826
```

```
RESULT 4
BM919548
LOCUS       BM919548                1000 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6761913 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748847
            5', mRNA sequence.
ACCESSION   BM919548
VERSION     BM919548.1  GI:19369927
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1000)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12777  row: e  column: 08
            High quality sequence stop: 643.
FEATURES             Location/Qualifiers
     source          1. .1000
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5748847"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_120"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           56.3%;  Score 726.8;  DB 3;  Length 1000;
  Best Local Similarity   95.9%;  Pred. No. 7e-186;
  Matches  800;  Conservative    0;  Mismatches   24;  Indels   10;  Gaps    5;

Qy          7 CGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTG 66
              ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CGATCAGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTG 61

Qy         67 GGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCC 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCC 121

Qy        127 TAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATG 181

Qy        187 GGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTG 241

Qy        247 GGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTC 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTC 301

Qy        307 TCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAG 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAG 361

Qy        367 CAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATG 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATG 421

Qy        427 AAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTG 486
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        422 AAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCAAATGCTG 481

Qy        487 ATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCC 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 ATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCC 541

Qy        547 CCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAG 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAG 601

Qy        607 GGAAGCACGTCTGTGGCCATCATCT-GCATACGGTGGGCGGTGTTGTCGAGAGGGATGTG 665
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        602 GGAAGCACGTCTGTGGCCATCATCTNGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTG 661

Qy        666 TGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAG 721

Qy        726 AGCAGACCACGGCGCC--AAGGCGAGGTCACGGTCCTTTCTGTTGGCAG-----ATTTAG 778
              |||||||||||||| | | |||||||||| |  |  | |  ||   |       ||
Db        722 AGCAGACCACGGCGNCCAAGGCCGAGGTCACCGGTCCCTTTTCTGTTGGCCAGATTTAGA 781

Qy        779 AGTTACAAAAGT-GGAGCACAAGTCAAACCAGAA-GGAACGGAGAAGCCTGATG 830
              |||||||||||| |||||||||||||||||||| |||||| |||||||| |||
Db        782 AGTTACAAAAGTGGGAGCACAAGTCAAACCAGAANGGAACGAAGAAGCCTNATG 835


RESULT 5
BQ807844
LOCUS       BQ807844                620 bp    mRNA    linear   EST 31-JUL-2002
DEFINITION  NISC_kk10h07.y1 NCI_CGAP_Brn72 Macaca mulatta cDNA clone
            IMAGE:5331301 5', mRNA sequence.
ACCESSION   BQ807844
VERSION     BQ807844.1  GI:22032053
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 620)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation:
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM11840  row: O  column: 14
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .620
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /db_xref="taxon:9544"
                     /clone="IMAGE:5331301"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NCI_CGAP_Brn72"
                     /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: NotI;
                     Site_2: EcoRV; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.2 kb. Constructed by Invitrogen.
                     Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match           46.0%;  Score 594.4;  DB 5;  Length 620;
  Best Local Similarity   97.4%;  Pred. No. 5.9e-150;
  Matches  604;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         26 GGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGA 85
```

```
                    |||||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db          1 GGAGCGGCTCTGCCTTCCCGAGCGCCGGACGCGGCGCCCTGGGGAAGGAGGGCGAAGCGA 60

Qy         86 CGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGT 145
                ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGCGATGGCTCTGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGT 120

Qy        146 GGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCA 205
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCA 180

Qy        206 CTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATA 265
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATA 240

Qy        266 TATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTG 325
                |||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        241 TATTGCATATGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTCGGCGTCCTCATTTG 300

Qy        326 CCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGA 385
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGA 360

Qy        386 GGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGG 445
                |||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||||
Db        361 AGAAAAGGTTGAAAAGATAGAATTGAACGACAGTATGAATGAAAACAGTGACACTGTTGG 420

Qy        446 GCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGT 505
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGT 480

Qy        506 AGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCC 565
                ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCGGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCC 540

Qy        566 GCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCA 625
                || |||||||||||||||||||||||||||||| || || |||||| ||||||||||
Db        541 ACCCGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACCCCGGGAAAGCACATCTGTGGCCA 600

Qy        626 TCATCTGCATACGGTGGGCG 645
                ||||||||| ||||||||||
Db        601 TCATCTGCACACGGTGGGCG 620


RESULT 6
BU150311
LOCUS       BU150311                874 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_8121018 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6178627 5', mRNA sequence.
ACCESSION   BU150311
VERSION     BU150311.1  GI:22663843
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 874)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13558  row: h  column: 20
            High quality sequence stop: 592.
FEATURES             Location/Qualifiers
     source          1. .874
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6178627"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match           46.0%;  Score 594.4;  DB 5;  Length 874;
  Best Local Similarity   98.9%;  Pred. No. 6.4e-150;
  Matches  609;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        179 GGACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAA 238
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GTACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAA 241

Qy        239 CGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTAT 298
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTAT 301

Qy        299 CATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTAC 358
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTAC 361

Qy        359 AACAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAG 418
                |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AACAGAAGCATAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAG 421

Qy        419 TGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGC 478
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGC 481

Qy        479 GAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCC 538
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCC 541

Qy        539 CGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGG 598
                |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        542 CGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGTGCCTTTGTCACCAGGGGG 601

Qy        599 GACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAG 658
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAG 661

Qy        659 GGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTC 718
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTC 721

Qy        719 CAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCA-CGGTCCTTTCTGTTGGCAGATTTA 777
                |||||||||||||||||||| |||||||||| |||||||||||||||||||||||||||
Db        722 CAGAGAGAGCAGACCACGGCGCCCAGGCGAGGTCACCGGTCCTTTCTGTTGGCAGATTTA 781

Qy        778 GAGTTACAAAAGTGGA 793
                ||||||| ||||| ||
Db        782 GAGTTACCAAAGTTGA 797


RESULT 7
DR004573
```

```
LOCUS       DR004573                 577 bp    mRNA    linear   EST 17-MAY-2005
DEFINITION  TC104761 Human placenta, large insert, pCMV expression library Homo
            sapiens cDNA clone TC104761 5' similar to Homo sapiens hypothetical
            protein LOC253981 (LOC253981), mRNA sequence.
ACCESSION   DR004573
VERSION     DR004573.1  GI:66264446
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
            Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
            Zhang,X., Jay,G. and He,W.
  TITLE     High-throughput cloning of full-length human cDNAs directly from
            cDNA libraries optimized for large and rare transcripts
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Kovacs, KF
            High Throughput cDNA Cloning
            OriGene Technologies, Inc. ( www.origene.com )
            6 Taft Court, Suite 100, Rockville, MD 20850, USA
            Tel: 301 340 3188
            Fax: 301 340 8606
            Email: cDNA@origene.com
            This EST submission is part of an on-going human full-length
            cloning project at OriGene Technologies, Inc.
            Please contact OriGene for access.
            OriGene Technologies, Inc.
            6 Taft Ct. Suite 100
            Rockville, MD 20850
            Tel: (301) 340-3188
            http://www.origene.com
            Seq primer: pCMV6 5prime forward vector primer, OriGene
            Technologies Inc.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TC104761"
                     /tissue_type="Placenta"
                     /clone_lib="Human placenta, large insert, pCMV expression
                     library"
                     /note="Vector: pCMV6-XL4; Site_1: EcoR1; Site_2: Xho1/Sal1
                     compatible end ligatio; Oligo-dT primed reverse
                     transcription optimized for large and GC rich mRNA
                     transcripts, cDNA size selection, optimized ligation for
                     large inserts into mammalian expression vector, random
                     clones selected for end sequence verification of
                     full-length genes"
ORIGIN

  Query Match             44.2%;  Score 570;  DB 8;  Length 577;
  Best Local Similarity   99.7%;  Pred. No. 2.5e-143;
  Matches  570;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         51 GGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTC 110
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GGGNCGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTC 65

Qy        111 CCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTG 170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTG 125

Qy        171 GTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCG 230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCG 185

Qy        231 CCCAGCAACGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTG 290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CCCAGCAACGATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTG 245

Qy        291 TTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTAT 350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTAT 305

Qy        351 CGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTG 410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTG 365

Qy        411 AATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAA 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 AATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAA 425

Qy        471 AATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCT 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 AATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCT 485

Qy        531 GAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCA 590
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 GANAGCCCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCA 545

Qy        591 CCAGGGGGGACGCCAGGGAAGCACGTCTGTGG 622
              ||||||||||||||||||||||||||||||||
Db        546 CCAGGGGGGACGCCAGGGAAGCACGTCTGTGG 577


RESULT 8
DN948289
LOCUS       DN948289                 815 bp    mRNA    linear   EST 29-APR-2005
DEFINITION  AGENCOURT_50711126 NCI_CGAP_Pr49 Rattus norvegicus cDNA clone
            IMAGE:7932431 5', mRNA sequence.
ACCESSION   DN948289
VERSION     DN948289.1  GI:62971409
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 815)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeff Green/Paturu Kondaiah, NCI.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM16355  row: k  column: 21
            High quality sequence stop: 792.
FEATURES             Location/Qualifiers
     source          1. .815
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="IMAGE:7932431"
                     /sex="male"
                     /tissue_type="ventral prostate, pool of 3-, 5-, and 7-days
                     post-castration"
                     /dev_stage="adult, 11 week"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Pr49"
                     /note="Organ: prostate; Vector: pCMV-SPORT6.1; Site_1:
                     NotI; Site_2: EcoRV; Cloned unidirectionally. Primer:
                     Oligo dT. Pool of 3 primary  libraries: NCI_CGAP_Pr30
                     (ventral prostate from 11 wk male, 3 days
```

post-castration, average insert size 2 kb), NCI_CGAP_Pr40 (ventral prostate from 11 wk male, 5 days post-castration, average insert size 1.6 kb) and NCI_CGAP_Pr41 (ventral prostate from 11 wk male, 7 days post-castration, average insert size 2.5 kb). Constructed by Life Technologies/Invitrogen. Note: this is a NCI_CGAP Library."

ORIGIN

```
  Query Match             42.9%;  Score 554.2;  DB 8;  Length 815;
  Best Local Similarity   80.9%;  Pred. No. 5.3e-139;
  Matches  659;  Conservative    0;  Mismatches  153;  Indels    3;  Gaps    1;

Qy         19 CGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGC 78
              | || |  ||||  | |   || || | || |||||||  |   || ||| ||| ||| |
Db          1 CAGCCCTCGAGCTTCACCCTCTACCTGGGCTCGGGACGTCGAAGCCGGGGCGAGAAGGAC 60

Qy         79 GAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTG 138
                || ||| ||||||||||||||| ||| ||||| |||||| || || || || | | | |
Db         61 AGAGGGACTCGGCGATGGCTCCGTGGGGACTCCTGGGGTCTGCTGTGCTTGCTGGTTCGG 120

Qy        139 TCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCA 198
              | |||||||||||||| |||||||||||||||||| || ||||||||     | |||||||
Db        121 TGTTCGTGGGAGGCGCTGTGAGTTCGCCGCTGGTAGCCCCGGACAACACTGGTAGCCGCA 180

Qy        199 CATTGCACTCCAGAACAGAGACGACCCCGTC---GCCCAGCAACGATACTGGGAATGGAC 255
              | || |||||  || ||||||| || |||||     |||| |||| |  ||||||||||||
Db        181 CGTTACACTCTCGAGCAGAGACAACGCCGTCATTACCCACCAACAACCCTGGGAATGGAC 240

Qy        256 ACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCG 315
              |||||||||| || |||||  | ||||||||||||||||||||  ||||||||||  |||| |
Db        241 ACCCAGAATACATCGCATATGTGCTTGTCCCTGTGTTCTTCGTCATGGGTCTTCTTGGTG 300

Qy        316 TCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAG 375
              ||||||| |||||| | |||||||||||||||||||| ||||||||||| || || | |
Db        301 TCCTCATCTGCCACTTACTTAAGAAGAAAGGCTATCGGTGTACAACAGAGGCGGAACAGG 360

Qy        376 ATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTG 435
              |  | |||||||||||||| |||||||||| |||||||||||| | ||||| |||||||
Db        361 AAGTTGAAGAGGAAAAGGTGGAAAAGATAGAGTTGAATGACAGTATAAATGAGAACAGTG 420

Qy        436 ACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAA 495
              |||| |||||||| || |||||||| || |||||||||||||| ||||| ||  | || |
Db        421 ACACCGTTGGGCAGATTGTCCACTATATTATGAAAAATGAAGCAAATGCAGACATTTTGA 480

Qy        496 AGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACAC 555
              | || |||||||| ||||||||| ||  ||||  ||||||||| || || || |||||||
Db        481 AAGCTATGGTAGCTGATAACAGCATGGTTGATGTTGAAAGCCCTGTCACTCCTAGCACAC 540

Qy        556 CAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACG 615
              |||||||||| || |||||||| |||||  |||||||||||  || ||||||||||| |
Db        541 CAGGGAGCCCACCTGTGAGTCCGGGGCCCCTGTCACCAGGGGCCACCCCAGGGAAGCATG 600

Qy        616 TCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGT 675
              |||| || || ||||| || || ||||| |||||||| ||  |||| ||||| || || |
Db        601 TCTGCGGACACCATCTTCACACAGTGGGAGGTGTTGTTGAACGGGACGTGTGCCAGCGCT 660

Qy        676 GTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCAC 735
              |||||||||||| |||||||| |||||  || || ||||||  ||| |||||||  | || |
Db        661 GTAGGCACAAGCGATGGCACTTCATAAAACCTACCAACAAGACCAAAGAGAGCCGGCCTC 720

Qy        736 GGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGC 795
              |||| ||||| |||||||| ||||| ||||| ||||| || ||||| ||||||||  ||||
Db        721 GGCGACAAGGGGAGGTCACTGTCCTCTCTGTGGGCAGGTTCAGAGTCACAAAAGTAGAGC 780

Qy        796 ACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATG 830
              |||||||||||||||||| || |||||| |||||
Db        781 ACAAGTCAAACCAGAAGGAGCGCAGAAGCTTGATG 815
```

RESULT 9
BP438847
LOCUS BP438847 684 bp mRNA linear EST 30-DEC-2003

DEFINITION BP438847 full-length enriched swine cDNA library, adult lung Sus scrofa cDNA clone LNG010079F11 5', mRNA sequence.
ACCESSION BP438847
VERSION BP438847.1 GI:40428914
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 684)
AUTHORS Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H., Okumura,N., Hamasima,N. and Awata,T.
TITLE PEDE (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries
JOURNAL Nucleic Acids Res. 32 (1), D484-D488 (2004)
PUBMED 14681463
COMMENT Contact: Hirohide Uenishi
Animal Genome Laboratory, Genome Research Department
National Institute of Agrobiological Sciences
2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
Tel: +81-29-838-8627
Fax: +81-29-838-8627
Email: huenishi@affrc.go.jp
EST project with full-length enriched cDNA libraries carried out in Animal Genome Research Program (Japan) by National Institute of Agrobiological Sciences and STAFF-Institute
Single pass sequencing of clones derived from oligo-capped cDNA library
Vector sequences were eliminated by RepeatMasker version 2002/07/13 and crossmatch version 0.990319
Low quality bases were trimmed based on the quality values.

FEATURES Location/Qualifiers
source 1. .684
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="LNG010079F11"
/tissue_type="lung"
/dev_stage="adult"
/clone_lib="full-length enriched swine cDNA library, adult lung"

ORIGIN

```
  Query Match             41.7%;  Score 538.2;  DB 3;  Length 684;
  Best Local Similarity   86.9%;  Pred. No. 1.1e-134;
  Matches  591;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy         38 CCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGC 97
              |||    |||| |||||||||||| ||| |||||||||||||||||| |||||| ||||||
Db          5 CCTATTGGAGCCCGGGACGCGGCACCCCGGGGGAGGAGGGCGAAGAGACGCGGGGATGGC 64

Qy         98 TCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGT 157
              || ||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         65 TCTGCGGGGACTCCCGGGGTCCGCCGTCCTCGCCGCTGCTGTCTTCGTGGGAGGCGCCGT 124

Qy        158 GAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGA 217
              |||||||| || || || |||||||||  ||||||||| ||| || ||| ||||  |
Db        125 GAGTTCGCCACTCGTAGCCCCGGACAATGGTGGCAGCCGCATATTACATTCCGGAACCAA 184

Qy        218 GACGACCCCGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATATTGCATACGC 277
              |||||| || || |||| ||| | |||||||||||||||||||||||| || || ||
Db        185 GACGACGCCCTCACCCACAAACAACACTGGGAATGGACACCCAGAATATATCGCGTATGC 244

Qy        278 GCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAA 337
              ||| ||||| |||||||| |||||||||||||||||||||||||| ||||||||||||||
Db        245 GCTCGTCCCCGTGTTCTTCATCATGGGTCTCTTTGGCGTCCTCATCTGCCACCTGCTTAA 304

Qy        338 GAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGA 397
              |||||||| |||||||||||||||||||| |||||||  | ||||||||||||||||
Db        305 GAAGAAAGGTTATCGTTGTACAACAGAAGCCGAGCAAGACGTGGAAGAGGAAAAGGTTGA 364

Qy        398 AAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCA 457
```

```
              |||||||||  ||||  |||||  ||||||||||| ||||||||  |||||  ||  ||  |||||
Db        365 AAAGATAGAGATGAACGACAGCGTGAATGAAAATAGTGACACCGTTGGACAGATTGTCCA 424

Qy        458 CTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAG 517
              ||||||||||||||||||||||  |||||  |||||  ||  |||||  |||||  ||||||||||
Db        425 CTACATCATGAAAAATGAAGCAAATGCCGATGTTTTGAAGGCAATGGTGGCAGATAACAG 484

Qy        518 CCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCC 577
              |||||||||||||||||||||  ||||||||||||||  ||  ||||||||||  |||  |  ||
Db        485 CCTGTATGATCCTGAAAGCCCTGTGACCCCCAGCACTCCTGGGAGCCCGCCCGTGGGCCC 544

Qy        578 TGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATAC 637
              |||||  ||||||||  ||  ||  ||  |||||  ||||||  ||||  |||||  ||||||||  ||
Db        545 CGGGCCCTTGTCACCGGGAGGCACCCCAGGAAAGCACATCTGCGGCCACCATCTGCACAC 604

Qy        638 GGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTT 697
              ||||||||||||  ||||  ||  ||||  ||||||||||||  |||  |||||||||||||||
Db        605 GGTGGGCGGTGCTGTCAAGCGGGACGTGTGTCATCGATGTANGCATAAGCGGTGGCACTT 664

Qy        698 TATAAAGCCCACTAACAAGT 717
              ||  |  ||  |  |||||||
Db        665 CATCAGACCTGCCAACAAGT 684


RESULT 10
BI691144
LOCUS       BI691144                906 bp    mRNA    linear   EST 18-SEP-2001
DEFINITION  603314531F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5354671 5',
            mRNA sequence.
ACCESSION   BI691144
VERSION     BI691144.1  GI:15653773
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11901  row: m  column: 08
            High quality sequence stop: 753.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5354671"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam6"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
ORIGIN

  Query Match             37.8%;  Score 488.6;  DB 3;  Length 906;
  Best Local Similarity   82.4%;  Pred. No. 3.6e-121;
  Matches  573;  Conservative    0;  Mismatches  119;  Indels    3;  Gaps    1;
```

```
Qy        156 GTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCCAGAACA 215
              |||||||||||||||||||  | |||||     | ||||  |||  ||  ||||||  ||  ||
Db          1 GTGAGTTCGCCGCTGGTGGCCGCCGACAACACTGGTAGCCACACGTTACACTCCCGAGCA 60

Qy        216 GAGACGACCC---CGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATATTGCA 272
              |||||  ||  |   ||||  ||||  ||||  |  ||||||||||||||||  |||||  ||  ||
Db         61 GAGACCACTCCATCGTCACCCACCAACAACCCTGGGAATGGACACCCGGAATACATCGCG 120

Qy        273 TACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCCACCTG 332
              ||||  |||  ||||||||||||||  ||||||||||  ||||  ||||||||  ||  |||  |
Db        121 TACGTGCTCGTCCCTGTGTTCTTCGTCATGGGTCTCCTTGGTGTCCTCATCTGTCACTTA 180

Qy        333 CTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGGAAAAG 392
              |||||||||||||||||  ||  ||||||||||  ||  ||  ||  ||||||||||||||
Db        181 CTTAAGAAGAAAGGCTACCGGTGTACAACAGAGGCGGAACAGGAAGTCGAAGAGGAAAAG 240

Qy        393 GTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGCAAATC 452
              ||  ||||||||||||  ||||||||||||  |  |||||  ||||||||||  ||  ||  ||  ||
Db        241 GTGGAAAAGATAGAGTTGAATGACAGTATAAATGAGAACAGTGACACGGTCGGCCAGATT 300

Qy        453 GTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGAT 512
              |||||  |||||  ||||||||||||||||||||  |||  ||||  ||  ||  ||||||||  |||
Db        301 GTCCAGTACATTATGAAAAATGAAGCGAATGCAGATATCTTGAAAGCTATGGTAGCTGAT 360

Qy        513 AACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTG 572
              ||||||  ||   |||  |||||||||  ||  ||  ||  |||||  ||||||||||  ||  |||
Db        361 AACAGCGTGGGCGATATTGAAAGCCCTGTCACTCCTAGCACGCCAGGGAGCCCACCTGTG 420

Qy        573 AGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTG 632
              ||||||||||||  ||||||||||||  ||  ||||||||||  ||||||||  |||||||||
Db        421 AGTCCTGGGCCCCTGTCACCAGGGGCCACCCCAGGGAAGCATGTCTGTGGTCATCATCTG 480

Qy        633 CATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGG 692
              ||  ||  |||||  |||||  ||  ||  ||||||||||  ||  ||  ||  ||||||||||  |||
Db        481 CACACAGTGGGAGGTGTCGTTGAACGGGATGTGTGCCAGCGATGCAGGCACAAGCGATGG 540

Qy        693 CACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTC 752
              |||||  |||||  |||||  ||||||  |||  ||||  |  |  ||||||||  ||||||||||||
Db        541 CACTTCATAAAACCCACCAACAAGACCAAAGAGGGACGGCCACGGCGACAAGGCGAGGTC 600

Qy        753 ACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAG 812
              ||  |||||  |||||  |||||  ||  |||||  ||  ||||||||||||||||||||  ||||||
Db        601 ACCGTCCTCTCTGTGGGCAGGTTCAGAGTCACGAAAGTGGAGCACAAGTCAAATCAGAAG 660

Qy        813 GAACGGAGAAGCCTGATGTCTGTTAGTGGGGCTGA 847
              ||  ||  |||||  ||||||||||  |||||  |||
Db        661 GAGCGCAGAAGTTTGATGTCTGTCAGTGGCATTGA 695


RESULT 11
DN382824
LOCUS       DN382824                626 bp    mRNA    linear   EST 07-MAR-2005
DEFINITION  LIB38534_030_A09_T7_1 LIB38534 Canis familiaris cDNA clone
            LIB38534_30_A09, mRNA sequence.
ACCESSION   DN382824
VERSION     DN382824.1  GI:60564044
KEYWORDS    EST.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 626)
  AUTHORS   Staten,N.R.
  TITLE     Direct Submission (Staten,N.R.)
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Nick Staten
            Tel: 636 247 6855
            Email: nicholas.r.staten@pfizer.com.
FEATURES             Location/Qualifiers
     source          1. .626
                     /organism="Canis familiaris"
```

```
                     /mol_type="mRNA"
                     /db_xref="taxon:9615"
                     /clone="LIB38534_30_A09"
                     /tissue_type="unknown"
                     /lab_host="DH10B"
                     /clone_lib="LIB38534"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI"
ORIGIN

  Query Match             37.1%;  Score 479.6;  DB 8;  Length 626;
  Best Local Similarity   86.2%;  Pred. No. 9.3e-119;
  Matches  530;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy          63 CCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCC 122
               ||  | ||||||||||||||  | ||||| ||||| ||| |||  ||| ||||||||||
Db          12 CCCCGCGGAGGAGGGCGAAGGAAGGCGGCCATGGCCCCGTGGGAGCTCTCGGGGTCCGCG 71

Qy         123 GTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGAC 182
               ||||| ||||| |||||||||||||||||| ||| ||||||||||||||||| ||||||
Db          72 GTCCTGGCCGCCGCTGTCTTCGTGGGAGGCACCGCGAGTTCGCCGCTGGTGGCCCCGGAC 131

Qy         183 AATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACGAT 242
               | |||  |||||| ||| |||||||||| ||||||||||||||  |||||  ||| ||
Db         132 ATTGGTGGCAGCCACACGTTGCACTCCAGGACAGAGACGACCCCGGAGCCCACTAACAAT 191

Qy         243 ACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATG 302
               |||||||||||||||||||||||||| || || || |||||||||||||||||| ||||||
Db         192 ACTGGGAATGGACACCCAGAATATATCGCGTATGCCCTTGTCCCTGTGTTCTTCATCATG 251

Qy         303 GGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACA 362
               |||||||||| |||||||| || |||||||||||||||| ||||||||||||||||||||
Db         252 GGTCTCTTTGGTGTCCTCATCTGTCACCTGCTTAAGAAGAAGGGCTATCGTTGTACAACA 311

Qy         363 GAAGCAGAGCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTG 422
               ||||| ||||||||  | |||||||| ||||||||||||||||| ||||||||||||||
Db         312 GAAGCCGAGCAAGAGGTGGAAGAGGAGAAGGTTGAAAAGATAGAGTTGAATGACAGTGTC 371

Qy         423 AATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAAT 482
               |||||||| |||||||| |||||||| || ||||| |||||||||||||||||||| ||
Db         372 AATGAAAATAGTGACACCGTTGGGCAGATTGTCCAGTACATCATGAAAAATGAAGCAAAC 431

Qy         483 GCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTG 542
               | |||||| |||||||| |||||||||||||||||||||||||    || ||||||||||| |||
Db         432 GTTGATGTTTTAAAGGCAATGGTAGCAGATAACAGCCTGGGGGACCCTGAAAGCCCTGTG 491

Qy         543 ACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACG 602
               || |||||||| || |||||||||| ||||| || ||||| ||||| ||||| || ||
Db         492 ACTCCCAGCACTCCTGGGAGCCCGCCTGTGAGCCCGGGGCCCTTGTCCCCAGGAGGCACC 551

Qy         603 CCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGAT 662
               ||||| || ||| ||| |||||| ||||| || || ||  | |||| ||| |||||||||
Db         552 CCAGGAAAACACATCTATGGCCACCATCTACACACAGTANGAGGTGCTGTTGAGAGGGAC 611

Qy         663 GTGTGTCATCGGTGT 677
               |||||||||||||||
Db         612 GTGTGTCATCGGTGT 626


RESULT 12
BI689796
LOCUS       BI689796                795 bp    mRNA    linear   EST 18-SEP-2001
DEFINITION  603316010F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5355962 5',
            mRNA sequence.
ACCESSION   BI689796
VERSION     BI689796.1  GI:15652425
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11905  row: c  column: 03
            High quality sequence stop: 717.
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5355962"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam6"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
ORIGIN

  Query Match             35.6%;  Score 460.2;  DB 3;  Length 795;
  Best Local Similarity   79.4%;  Pred. No. 1.8e-113;
  Matches  620;  Conservative    0;  Mismatches  153;  Indels    8;  Gaps    6;

Qy          92 GATGGCTCCGCGGGCACTCCC-GGGGTCCGCCGTCCTAGCCGCTGCTGTCTTCGTGGGAG 150
               |||||||| | ||| |||||| |||| |||| || || || ||| | || ||||||||||
Db          17 GATGGCTCTGTGGGGACTCCCGGGGGGCCGCTGTGCTTGCTGCTTCAGTGTTCGTGGGAG 76

Qy         151 GCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGCACATTGCACTCCA 210
               |||| ||||||||||||||||||||  ||   |    |    |    |    |  ||||
Db          77 GCGCTGTGAGTTCGCCGCTGGTGGCCGCGACAACACTGGTAGCCACACGTTACACCTCCC 136

Qy         211 GAACAGAGACGACCC---CGTCGCCCAGCAACGATACTGGGAATGGACACCCAGAATATA 267
               || ||||||| || |   |||| |||| |||| |  ||||||||||||||||| ||||| |
Db         137 GAGCAGAGACCACTCCATCGTCACCCACCAACAACCCTGGGAATGGACACCCGGAATACA 196

Qy         268 TTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCATTTGCC 327
               | || |||| ||| ||||||||||||||  |||||||||| |||| |||||||| || |
Db         197 TCGCGTACGTGCTCGTCCCTGTGTTCTTCGTCATGGGTCTCCTTGGTGTCCTCATCTGTC 256

Qy         328 ACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGAAGAGG 387
               || | ||||||||||||||||||| || |||||||||| || || || ||  |||||||||
Db         257 ACTTACTTAAGAAGAAAGGCTACCGGTGTACAACAGAGGCGGAACAGGAAGTCGAAGAGG 316

Qy         388 AAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGTTGGGC 447
               ||||||| |||||||||| |||||||||||| | ||||| |||||||||| || || |
Db         317 AAAAGGTGGAAAAGATAGAGTTGAATGACAGTATAAATGAGAACAGTGACACGGTCGGCC 376

Qy         448 AAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGATGGTAG 507
               | || ||||| ||||| |||||||||||||||||||| ||| |||| || || |||||||
Db         377 AGATTGTCCAGTACATTATGAAAAATGAAGCGAATGCAGATATCTTGAAAGCTATGGTAG 436

Qy         508 CAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAGCACACCAGGGAGCCCGC 567
               | |||||||||  ||   ||| |||||||||| || || ||||| |||||||||||| |
Db         437 CTGATAACAGCGTGGGCGATTTTGAAAGCCCTGTCACTCCTAGCACGCCAGGGAGCCCAC 496

Qy         568 CAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAAGCACGTCTGTGGCCATC 627
               | |||||||||||||  |||||||| |||  || ||||| |||| |||||||| ||||
Db         497 CTGTGAGTCCTGGGCCCCTGTCACCA-GGGCCACCCCAGGCAAGCATGTCTGTGGTCATC 555

Qy         628 ATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCATCGGTGTAGGC-ACAAG 686
               ||||||| || ||||| ||||  || ||  |||||||||| || || || |||| |||||
```

```
Db        556 ATCTGCACACAGTGGGAGGTGGCGTTGAACGGGATGTGTGCCAGCGATGCAGGCAACAAG 615

Qy        687 CGGTGGCACTTTATAAAGCCCACTAACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGC 746
              || |||||||| ||||| ||||| | |||| ||| |||| | |   ||||||| ||||||
Db        616 CGATGGCACTTCATAAAACCCACCATCAAGACCACAGAG-GGACGGCACGGCGACAAGGC 674

Qy        747 GAGGTCACGGTCCTTTCTGTTGGCAGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAAC 806
              |||||||| ||||| ||||| ||||| || ||||| || ||||||||||||||||||||
Db        675 GAGGTCACCGTCCTCTCTGTGGGCAGGTTCAGAGTCACGAAAGTGGAGCACAAGTCAAAT 734

Qy        807 CAGAAGGAACGGAGAAGCCTGATG-TCTGTTAGTGGGGCTGAAACCGTCAATGGGGAGGT 865
              |||||||| || ||||   ||||| ||||| ||||||  |  ||  ||||  ||||| ||
Db        735 CAGAAGGAGCGCAGAACTTTGATGTTCTGTCAGTGGGCATTGAAGTGTCACCGGGGATGT 794

Qy        866 G 866
              |
Db        795 G 795
```

RESULT 13
BY754410
LOCUS BY754410 680 bp mRNA linear EST 17-DEC-2002
DEFINITION BY754410 RIKEN full-length enriched, osteoclast-like cell Mus musculus cDNA clone I420039H17 5', mRNA sequence.
ACCESSION BY754410
VERSION BY754410.1 GI:27186339
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 680)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Takashi Ishikawa (Department of Surgery 2 Yokohama City University 3-9 Fukuura,Kanazawa-ku,Yokohama 236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .680
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I420039H17"
/cell_type="osteoclast-like cell"
/clone_lib="RIKEN full-length enriched, osteoclast-like cell"
ORIGIN

Query Match 35.2%; Score 454.6; DB 5; Length 680;
Best Local Similarity 81.0%; Pred. No. 5.9e-112;
Matches 542; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

```
Qy         14 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGG 73
              |||| | || |  |||| || | | || |||   | |||||||| ||| || ||| ||||
Db         12 CCCGGCAGCCCTCGAGCTGCACCGTCTACCCCGACCCGGGACGCCGCGGCCGGGGCGAGG 71

Qy         74 AGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGC 133
              |||||  || |||||||||||||||||| | ||| |||||||||||||||| || || || ||
Db         72 AGGGCAGAGGGACGCGGCGATGGCTCTGTGGGGACTCCCGGGGTCCGCTGTGCTTGCTGC 131

Qy        134 TGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAG 193
              | | || |||||||||||||| |||||||||||||||||||  | |||||     | ||
Db        132 TTCAGTGTTCGTGGGAGGCGCTGTGAGTTCGCCGCTGGTGGCCGCCGACAACACTGGTAG 191

Qy        194 CCGCACATTGCACTCCAGAACAGAGACGACCC---CGTCGCCCAGCAACGATACTGGGAA 250
              || ||| || |||||| || ||||||| || |   |||| |||| |||| |  |||||||
Db        192 CCACACGTTACACTCCCGAGCAGAGACCACTCCATCGTCACCCACCAACAACCCTGGGAA 251

Qy        251 TGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTT 310
              |||||||| ||||| || || |||| ||| ||||||||||||||  |||||||||| |
Db        252 TGGACACCCGGAATACATCGCGTACGTGCTCGTCCCTGTGTTCTTCGTCATGGGTCTCCT 311

Qy        311 TGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGA 370
              ||| |||||||| || ||| | |||||||||||||||| || |||||||||| || ||
Db        312 TGGTGTCCTCATCTGTCACTTACTTAAGAAGAAAGGCTACCGGTGTACAACAGAGGCGGA 371

Qy        371 GCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAA 430
                || ||  |||||||||||||||| |||||||||| |||||||||||| | ||||| ||
Db        372 ACAGGAAGTCGAAGAGGAAAAGGTGGAAAAGATAGAGTTGAATGACAGTATAAATGAGAA 431

Qy        431 CAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGT 490
```

```
Db        432 CAGTGACACGGTCGGCCAGATTGTCCAGTACATTATGAAAAATGAAGCGAATGCAGATAT 491

Qy        491 CTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAG 550

Db        492 CTTGAAAGCTATGGTAGCTGATAACAGCGTGGGCGATATTGAAAGCCCTGTCACTCCTAG 551

Qy        551 CACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAA 610

Db        552 CACGCCAGGGAGCCCACCTGTGAGTCCTGGGCCCCTGTCACCAGGGGCCACCCCAGGGAA 611

Qy        611 GCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCA 670

Db        612 GCATGTCTGTGGTCATCATCTGCACACAGTGGGAGGTGTCGTTGACGGGGATGTGTGCCA 671

Qy        671 TCGGTGTAG 679

Db        672 GCGATGCAG 680


RESULT 14
BI658107
LOCUS       BI658107                757 bp    mRNA    linear   EST 12-SEP-2001
DEFINITION  603285885F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5330037 5',
            mRNA sequence.
ACCESSION   BI658107
VERSION     BI658107.1  GI:15572343
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11837  row: j  column: 22
            High quality sequence stop: 741.
FEATURES             Location/Qualifiers
     source          1. .757
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="NMRI"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5330037"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam4"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Priscilla Furth,
                     NIH Reference for transgenic model: Li et al., Cell Growth
                     and  Differentiation 7, 3-11 (1996)."
ORIGIN

  Query Match            35.2%;  Score 454.2;  DB 3;  Length 757;
  Best Local Similarity  80.0%;  Pred. No. 7.7e-112;
  Matches  572;  Conservative    0;  Mismatches  138;  Indels    5;  Gaps    3;

Qy          3 CACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGC 62

Db         36 CATTCGGTCCACCCGGCAGCCCTCGAGCTGCACCGTCTACCCCGACCCGGGACGCCGCGG 95

Qy         63 CCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCC 122

Db         96 CCGGGGCGAGGAGGGCAGAGGGACGCGGCGATGGCTCTGTGGGGACTCCCGGGGTCCGCT 155

Qy        123 GTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGAC 182

Db        156 GTGCTTGCTGCTTCAGTGTTCGTGGGAGGCGCTGTGAGTTCGCCGCTGGTGGCCGCCGAC 215

Qy        183 AATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCC---CGTCGCCCAGCAAC 239

Db        216 AACACTGGTAGCCACACGTTACACTCCCGAGCAGAGACCACTCCATCGTCACCCACCAAC 275

Qy        240 GATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATC 299

Db        276 AACCCTGGGAATGGACACCCGGAATACATCGCGTACGTGCCCGTCCCTGTGTTCTTCGTC 335

Qy        300 ATGGGTCTCTTTGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACA 359

Db        336 ATGGGTCTCCTTGGTGTCCTCATCTGTCACTTACTTAAGAAGAAAGGCTACCGGTGTACA 395

Qy        360 ACAGAAGCAGAGCAAGATATCGAAG-AGGAAAAGGTTGAAAAGATAGAATTGAATGACAG 418

Db        396 ACAGAGGCGGAACAGGAAGTCGAAGCAGGACAAGGTGGCAAAGATAGAGTTGAATGACAG 455

Qy        419 TGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGC 478

Db        456 TATAAATGAGAACAGTGACACGGTCGGCCAGATTGTCCAGTACATTATGAAAAATGAAGC 515

Qy        479 GAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCC 538

Db        516 GAATGCAGATATCTTGAAAGCTATGGTAGCTGATAACAGCGTGGGCGATATTGAAAGCCC 575

Qy        539 CGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGG 598

Db        576 TGTCACTCCTAGCACGCCAGGGAGCCCACCTGTGAGTCCTGGGCCCCTGTCACCAGGGGC 635

Qy        599 GACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAG 658

Db        636 CACCCCAGGGAAGCATGTCTGTGGTCATCATCTGCACACAGTGGGAGGTGTCGTTGAACG 695

Qy        659 GGATGTG-TGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAA 712

Db        696 GGATGTGTTGCCAGCGATGCAGGCACAAGCGATGGCACTTCATAAAAACCACAAA 750


RESULT 15
BY754598
LOCUS       BY754598                672 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY754598 RIKEN full-length enriched, osteoclast-like cell Mus
            musculus cDNA clone I420048H13 5', mRNA sequence.
ACCESSION   BY754598
VERSION     BY754598.1  GI:27186642
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 672)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
```

Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Takashi Ishikawa (Department of Surgery 2 Yokohama City University 3-9 Fukuura,Kanazawa-ku,Yokohama 236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers

source 1. .672
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I420048H13"
/cell_type="osteoclast-like cell"
/clone_lib="RIKEN full-length enriched, osteoclast-like cell"

ORIGIN

Query Match 34.8%; Score 449.4; DB 5; Length 672;
Best Local Similarity 81.1%; Pred. No. 1.5e-110;
Matches 535; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

```
Qy      14 CCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCCCTGGGGGAGG 73
           |||| | || |  |||| || | | || |||   | ||||||||| ||| || ||| ||||
Db      12 CCCGGCAGCCCTCGAGCTGCACCGTCTACCCCGACCCGGGACGCCGCGGCCGGGGCGAGG 71

Qy      74 AGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGC 133
           |||||  || ||||||||||||||| | ||| |||||||||||||| || || || ||
Db      72 AGGGCAGAGGGACGCGGCGATGGCTCTGTGGGGACTCCCGGGGTCCGCTGTGCTTGCTGC 131

Qy     134 TGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAG 193
           | | || |||||||||||||| |||||||||||||||||||  | |||||     | ||
Db     132 TTCAGTGTTCGTGGGAGGCGCTGTGAGTTCGCCGCTGGTGGCCGCCGACAACACTGGTAG 191

Qy     194 CCGCACATTGCACTCCAGAACAGAGACGACCC---CGTCGCCCAGCAACGATACTGGGAA 250
           || ||| || |||||| || |||||| || |   |||| |||| |||| |  |||||||
Db     192 CCACACGTTACACTCCCGAGCAGAGACCACTCCATCGTCACCCACCAACAACCCTGGGAA 251

Qy     251 TGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTT 310
           |||||||| ||||| || || |||| ||| |||||||||||||  |||||||||| |
Db     252 TGGACACCCGGAATACATCGCGTACGTGCTCGTCCCTGTGTTCTTCGTCATGGGTCTCCT 311

Qy     311 TGGCGTCCTCATTTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGA 370
           ||| |||||||| || ||| | |||||||||||||||||| || |||||||||| || ||
Db     312 TGGTGTCCTCATCTGTCACTTACTTAAGAAGAAAGGCTACCGGTGTACAACAGAGGCGGA 371

Qy     371 GCAAGATATCGAAGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAA 430
           || ||  |||||||||||||||| |||||||||| |||||||||| | ||||| ||
Db     372 ACAGGAAGTCGAAGAGGAAAAGGTGGAAAAGATAGAGTTGAATGACAGTATAAATGAGAA 431

Qy     431 CAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGT 490
           ||||||||| || || || || ||||| ||||| |||||||||||||||||||| ||| |
Db     432 CAGTGACACGGTCGGCCAGATTGTCCAGTACATTATGAAAAATGAAGCGAATGCAGATAT 491

Qy     491 CTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGCCCCGTGACCCCCAG 550
           ||| || || |||||||| ||||||||| ||   |||  ||||||||| || || || ||
Db     492 CTTGAAAGCTATGGTAGCTGATAACAGCGTGGGCGATATTGAAAGCCCTGTCACTCCTAG 551

Qy     551 CACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGGGGGACGCCAGGGAA 610
           ||| |||||||||| || |||||||||||||  |||||||||  ||  || ||||||||
Db     552 CACGCCAGGGAGCCCACCTGTGAGTCCTGGGCCCCTGTCACCAGNGGCCACCCCAGGGAA 611

Qy     611 GCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAGAGGGATGTGTGTCA 670
           ||| |||||||| |||||||||| || ||||| ||||| || ||  |||||||||| ||
Db     612 GCATGTCTGTGGTCATCATCTGCACACAGTGGGAGGTGTCGTTGAACGGGATGTGTGCCA 671
```

Search completed: February 23, 2006, 16:10:55
Job time : 4940 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: February 23, 2006, 14:58:13 ; Search time 269 Seconds
(without alignments)
8530.974 Million cell updates/sec

Title: US-10-620-562-1
Perfect score: 1291
Sequence: 1 cccacgcgtccggcccgccg.....cttacgcaaaaaaaaaaaaa 1291

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1196.6	92.7	1603	3	US-09-205-258-56	Sequence 56, Appl
	2	1196.6	92.7	1603	3	US-10-004-860-56	Sequence 56, Appl
	3	72.6	5.6	1558	3	US-09-949-016-5115	Sequence 5115, Ap
	4	72.6	5.6	1793	3	US-09-774-528-2	Sequence 2, Appli
	5	72.6	5.6	1793	3	US-10-120-988-2	Sequence 2, Appli
c	6	63.6	4.9	7218	2	US-08-232-463-14	Sequence 14, Appl
	7	51.4	4.0	16102	3	US-09-949-016-16857	Sequence 16857, A
c	8	51.4	4.0	19927	3	US-09-949-016-16044	Sequence 16044, A
	9	45	3.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	10	44.6	3.5	433	3	US-09-621-976-10505	Sequence 10505, A
	11	44.4	3.4	573	3	US-09-902-540-8667	Sequence 8667, Ap
	12	44.4	3.4	9191	3	US-09-902-540-918	Sequence 918, App
	13	43	3.3	4563	3	US-10-012-231A-259	Sequence 259, App
	14	43	3.3	4563	3	US-10-015-389A-259	Sequence 259, App
	15	43	3.3	4563	3	US-10-006-768A-259	Sequence 259, App
	16	43	3.3	4563	3	US-10-015-671A-259	Sequence 259, App
	17	43	3.3	4563	3	US-10-015-393A-259	Sequence 259, App
	18	43	3.3	4563	3	US-10-011-833A-259	Sequence 259, App
	19	43	3.3	4563	3	US-10-006-041A-259	Sequence 259, App
	20	43	3.3	4563	3	US-10-012-064A-259	Sequence 259, App
c	21	42.2	3.3	58857	3	US-09-477-962-1	Sequence 1, Appli
	22	42	3.3	50229	3	US-09-949-016-15499	Sequence 15499, A
	23	41.6	3.2	505	3	US-09-621-976-15639	Sequence 15639, A
c	24	41	3.2	1074	3	US-09-252-991A-15993	Sequence 15993, A
	25	41	3.2	1287	3	US-09-252-991A-16510	Sequence 16510, A
c	26	41	3.2	1356	3	US-09-252-991A-16093	Sequence 16093, A
	27	40.6	3.1	663	3	US-09-489-039A-6801	Sequence 6801, Ap
c	28	40.4	3.1	8625	3	US-09-949-016-16566	Sequence 16566, A
c	29	40.4	3.1	15632	3	US-09-949-016-15119	Sequence 15119, A
	30	40	3.1	2988	3	US-09-949-016-2383	Sequence 2383, Ap
	31	40	3.1	5075	3	US-09-949-016-100	Sequence 100, App
	32	40	3.1	5089	9	5177197-31	Patent No. 5177197
c	33	39.8	3.1	4010	2	US-08-785-310A-3	Sequence 3, Appli
c	34	39.8	3.1	140844	3	US-09-949-016-14199	Sequence 14199, A
c	35	39.6	3.1	6090	3	US-09-631-603-1	Sequence 1, Appli
	36	39.4	3.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	37	39	3.0	24043	3	US-09-949-016-16104	Sequence 16104, A
c	38	38.6	3.0	49301	3	US-09-949-016-16296	Sequence 16296, A
	39	38.4	3.0	852	3	US-09-902-540-5894	Sequence 5894, Ap
	40	38.4	3.0	1403	3	US-09-902-540-163	Sequence 163, App
c	41	38.4	3.0	1471	3	US-09-949-016-5372	Sequence 5372, Ap
	42	38.4	3.0	16340	3	US-09-949-002-583	Sequence 583, App
	43	38.4	3.0	16341	3	US-09-949-016-14579	Sequence 14579, A
	44	38.4	3.0	16341	3	US-09-949-016-14580	Sequence 14580, A
	45	38.4	3.0	16341	3	US-09-949-002-747	Sequence 747, App

ALIGNMENTS

RESULT 1
US-09-205-258-56
; Sequence 56, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892

```
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
;   LENGTH: 1603
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (328)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (336)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (341)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (788)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-56

  Query Match            92.7%;  Score 1196.6;  DB 3;  Length 1603;
  Best Local Similarity  97.9%;  Pred. No. 0;
  Matches 1268;  Conservative    5;  Mismatches   13;  Indels    9;  Gaps    6;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||

Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||| |  :     | ||||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||||| | ||||||||||||:|||:||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141
```

```
Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
             |||||||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296
```

RESULT 2
US-10-004-860-56
; Sequence 56, Application US/10004860
; Patent No. 6914047
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/10/004,860
; CURRENT FILING DATE: 2001-12-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 1603
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (328)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (336)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (341)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (788)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-004-860-56

Query Match 92.7%; Score 1196.6; DB 3; Length 1603;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1268; Conservative 5; Mismatches 13; Indels 9; Gaps 6;

```
Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  | ||
Db          7 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGA--CCGC 64

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         65 GCCCTGGGGGAGGAGGGCGAA-CGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 123

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 183

Qy        181 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCCAGCAACG 243

Qy        241 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATACTGGGAATGGACACCCAGAATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCA 303

Qy        301 TGGGTCTCTTTGGCGTCCTCATTTGCCACCTGC---TTAAGAAGAAAGGCTATCGTTGTA 357
              |||||||||||||||||||||||| |  :    | ||||||||||||||||||||||||
Db        304 TGGGTCTCTTTGGCGTCCTCATTTNGCCAMCTNGCTTNAAGAAGAAAGGCTATCGTTGTA 363

Qy        358 CAACAGAAGCAGAGCAAGATATCGAAGAGG-AAAAGGTTGAAAAGATAGAATTGAATGAC 416
              |||||||||||||||||||||||||||| | |||||||||||||:|||:||||||||||
Db        364 CAACAGAAGCAGAGCAAGATATCGAAGAAGAAAAAGGTTGAAAAGWTAGRATTGAATGAC 423

Qy        417 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTGTGAATGAAAACAGTGACACTGTTGGGCAAATCGTCCACTACATCATGAAAAATGAA 483

Qy        477 GCGAATGCTGATGTCTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 536
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGAATGCTGATGTYTTAAAGGCGATGGTAGCAGATAACAGCCTGTATGATCCTGAAAGC 543

Qy        537 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGGCCTTTGTCACCAGGG 596
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        544 CCCGTGACCCCCAGCACACCAGGGAGCCCGCCAGTGAGTCCTGGG-CTTTGTCACCAGGG 602

Qy        597 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTTGTCGAG 656
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        603 GGGACGCCAGGGAAGCACGTCTGTGGCCATCATCTGCATACGGTGGGCGGTGTWGTCGAG 662

Qy        657 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGGGATGTGTGTCATCGGTGTAGGCACAAGCGGTGGCACTTTATAAAGCCCACTAACAAG 722

Qy        717 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGCAGATTT 782

Qy        777 AGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 836
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGAGTNACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATGTCTGTT 842

Qy        837 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAACGCAGT 902

Qy        897 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTGCAGGGT 962

Qy        957 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATTTGACAG 1022

Qy       1017 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 GGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGATTACTTG 1082

Qy       1077 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGGCCTGT 1136
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1083 TCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTGCCTCCAGGGGCTTGG-CTGT 1141

Qy       1137 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GTGATAATGAGGATAGAGGATTACTTGTGAGGCAATGTGGCATGGTGGGGATTGTGGCAA 1201

Qy       1197 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 ACTAGAATTCACATCACCCACCATATAGGGCTTGCATTACCACGAGGCAGAAAGCACCTA 1261

Qy       1257 GTGTTGCTGCATCTTCTTACGCAAAAAAAAAAAAA 1291
              |||||||||||||||||||||||||||||| | ||||
Db       1262 GTGTTGCTGCATCTTCTTACGCAAAAAAGACAAAA 1296
```

RESULT 3
US-09-949-016-5115
; Sequence 5115, Application US/09949016
; Patent No. 6812339

```
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5115
;   LENGTH: 1558
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-5115

  Query Match             5.6%;  Score 72.6;  DB 3;  Length 1558;
  Best Local Similarity   59.0%;  Pred. No. 1.5e-10;
  Matches  148;  Conservative    0;  Mismatches   94;  Indels    9;  Gaps    1;

Qy        263 ATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCAT 322
              ||||||    | |  |||||| |  || |||||  ||||||| ||  | ||| || | ||
Db        803 ATATATGCTCTTCCTGCTTGTGCTGGTCTTCTTCCTCATGGGCCTGGTAGGCTTCATGAT 862

Qy        323 TTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGA 382
              ||||||  |||| |||||||| |||||  || ||    ||       |    ||    ||
Db        863 CTGCCACGTGCTCAAGAAGAAGGGCTACCGCTGCCGCACGTCGAGGGGCTCTGAGCCTGA 922

Qy        383 AGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGT 442
              || |   || || |        ||      ||||||           ||    ||||| ||
Db        923 CGATGCCCAGCTTCAGCCCCCTGAGGACGATGACATG---------AATGAGGACACAGT 973

Qy        443 TGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGAT 502
              | |   || || | || ||||||  | ||||||| |||||||| | ||| |||| |||
Db        974 AGAGAGGATTGTTCGCTGCATCATCCAGAATGAAGCCAATGCTGAGGCCTTGAAGGAGAT 1033

Qy        503 GGTAGCAGATA 513
              | | |  || |
Db       1034 GCTGGGGGACA 1044


RESULT 4
US-09-774-528-2
; Sequence 2, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION:  No. 6743619el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE:  802
;  CURRENT APPLICATION NUMBER: US/09/774,528
;  CURRENT FILING DATE:  2001-01-30
;  NUMBER OF SEQ ID NOS:  441
;  SOFTWARE:  pt_FL_genes Version 2.0
; SEQ ID NO 2
;   LENGTH: 1793
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (756)..(1667)
US-09-774-528-2

  Query Match             5.6%;  Score 72.6;  DB 3;  Length 1793;
  Best Local Similarity   59.0%;  Pred. No. 1.7e-10;
  Matches  148;  Conservative    0;  Mismatches   94;  Indels    9;  Gaps    1;

Qy        263 ATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCAT 322
              ||||||    | |  |||||| |  || |||||  ||||||| ||  | ||| || | ||
Db        800 ATATATGCTCTTCCTGCTTGTGCTGGTCTTCTTCCTCATGGGCCTGGTAGGCTTCATGAT 859

Qy        323 TTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGA 382
              ||||||  |||| |||||||| |||||  || ||    ||       |    ||    ||
Db        860 CTGCCACGTGCTCAAGAAGAAGGGCTACCGCTGCCGCACGTCGAGGGGCTCTGAGCCTGA 919

Qy        383 AGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGT 442
              || |   || || |        ||      ||||||           ||    ||||| ||
Db        920 CGATGCCCAGCTTCAGCCCCCTGAGGACGATGACATG---------AATGAGGACACAGT 970

Qy        443 TGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGAT 502
              | |   || || | || ||||||  | ||||||| |||||||| | ||| |||| |||
Db        971 AGAGAGGATTGTTCGCTGCATCATCCAGAATGAAGCCAATGCTGAGGCCTTGAAGGAGAT 1030

Qy        503 GGTAGCAGATA 513
              | | |  || |
Db       1031 GCTGGGGGACA 1041


RESULT 5
US-10-120-988-2
; Sequence 2, Application US/10120988
; Patent No. 6919193
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Ren, Peiyan
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6919193el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 802CON
;  CURRENT APPLICATION NUMBER: US/10/120,988
;  CURRENT FILING DATE:  2002-04-11
;  PRIOR APPLICATION NUMBER: 09/774,528
;  PRIOR FILING DATE: 2001-01-30
;  NUMBER OF SEQ ID NOS: 441
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 2
;   LENGTH: 1793
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (756)..(1667)
US-10-120-988-2

  Query Match             5.6%;  Score 72.6;  DB 3;  Length 1793;
  Best Local Similarity   59.0%;  Pred. No. 1.7e-10;
  Matches  148;  Conservative    0;  Mismatches   94;  Indels    9;  Gaps    1;

Qy        263 ATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCAT 322
              ||||||    | |  |||||| |  || |||||  ||||||| ||  | ||| || | ||
Db        800 ATATATGCTCTTCCTGCTTGTGCTGGTCTTCTTCCTCATGGGCCTGGTAGGCTTCATGAT 859

Qy        323 TTGCCACCTGCTTAAGAAGAAAGGCTATCGTTGTACAACAGAAGCAGAGCAAGATATCGA 382
              ||||||  |||| |||||||| |||||  || ||    ||       |    ||    ||
Db        860 CTGCCACGTGCTCAAGAAGAAGGGCTACCGCTGCCGCACGTCGAGGGGCTCTGAGCCTGA 919
```

```
Qy        383 AGAGGAAAAGGTTGAAAAGATAGAATTGAATGACAGTGTGAATGAAAACAGTGACACTGT 442
              || |   || || |     ||     ||||||          ||    ||||| ||
Db        920 CGATGCCCAGCTTCAGCCCCCTGAGGACGATGACATG---------AATGAGGACACAGT 970

Qy        443 TGGGCAAATCGTCCACTACATCATGAAAAATGAAGCGAATGCTGATGTCTTAAAGGCGAT 502
              | |   || || | || |||||| | |||||||| |||||||| | ||| |||| |||
Db        971 AGAGAGGATTGTTCGCTGCATCATCCAGAATGAAGCCAATGCTGAGGCCTTGAAGGAGAT 1030

Qy        503 GGTAGCAGATA 513
              | | |  || |
Db       1031 GCTGGGGGACA 1041


RESULT 6
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             4.9%;  Score 63.6;  DB 2;  Length 7218;
  Best Local Similarity   5.2%;  Pred. No. 1.5e-07;
  Matches   21;  Conservative  228;  Mismatches  157;  Indels    0;  Gaps    0;

Qy        711 AACAAGTCCAGAGAGAGCAGACCACGGCGCCAAGGCGAGGTCACGGTCCTTTCTGTTGGC 770
              :: :::   :::::::: ::::  : :: :  :::: ::::  : ::        :  ::
Db       1425 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1366

Qy        771 AGATTTAGAGTTACAAAAGTGGAGCACAAGTCAAACCAGAAGGAACGGAGAAGCCTGATG 830
              :::   ::::  : ::::: :::: : :::  :::  :::::::: :::::::   :: :
Db       1365 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1306

Qy        831 TCTGTTAGTGGGGCTGAAACCGTCAATGGGGAGGTGCCGGCAACACCTGTGAAGAGAGAA 890
              :  :: :::: ::::  :  :: ::::::: :  :: :: :   : ::::::::::
Db       1305 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1246

Qy        891 CGCAGTGGCACAGAGTAGCAGGTGAGCCGTGGTTTTGGTGACATTGGGGGCAGAGTGGTG 950
              : :: :: : :::: :: ::: :::  : ::    :: :: :  ::::: :::: :: :
Db       1245 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1186

Qy        951 CAGGGTGAGGAGAAGGTACTTGGAGCCTCCCAGGTGCTGTGGCAGCATAGGAATGGTATT 1010
              :::: :::::::::: :   ::::      ::: :  : :: :: : ::::: :: :
Db       1185 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1126

Qy       1011 TGACAGGGAAGTGGGAGAGCTTTCCTTGACCCAGGAAGACTGAGGGGGACTGAACATGAT 1070
              :: ::::::: :::::::        ::   :::::::  ::::::::  ::: : ::
Db       1125 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRA 1066

Qy       1071 TACTTGTCTGCCTAGAGCTTCTTGTAAAGAAGTCACAAACTTAGTG 1116
              |       || ||| || || |    ||    |  |  || |  |||
Db       1065 TCGCAAGCTCCCTCGACCTGCAGCCAAGCTCGGAATTAATTCTGTG 1020


RESULT 7
US-09-949-016-16857
; Sequence 16857, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16857
;   LENGTH: 16102
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16857

  Query Match             4.0%;  Score 51.4;  DB 3;  Length 16102;
  Best Local Similarity   72.0%;  Pred. No. 0.00087;
  Matches   67;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy        263 ATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCAT 322
              ||||||     | |  |||||| |  || |||||  ||||||| ||   | ||| || | ||
Db      10563 ATATATGCTCTTCCTGCTTGTGCTGGTCTTCTTCCTCATGGGCCTGGTAGGCTTCATGAT 10622

Qy        323 TTGCCACCTGCTTAAGAAGAAAGGCTATCGTTG 355
              |||||| |||| |||||||| ||||| || ||
Db      10623 CTGCCACGTGCTCAAGAAGAAGGGCTACCGCTG 10655


RESULT 8
US-09-949-016-16044/c
; Sequence 16044, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
```

```
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16044
;   LENGTH: 19927
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16044

  Query Match             4.0%;  Score 51.4;  DB 3;  Length 19927;
  Best Local Similarity   72.0%;  Pred. No. 0.00098;
  Matches   67;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy        263 ATATATTGCATACGCGCTTGTCCCTGTGTTCTTTATCATGGGTCTCTTTGGCGTCCTCAT 322
              ||||||    | |  |||||| |  || |||||  ||||||| ||  | ||| || | ||
Db        532 ATATATGCTCTTCCTGCTTGTGCTGGTCTTCTTCCTCATGGGCCTGGTAGGCTTCATGAT 473

Qy        323 TTGCCACCTGCTTAAGAAGAAAGGCTATCGTTG 355
              |||||| |||| |||||||| ||||| || ||
Db        472 CTGCCACGTGCTCAAGAAGAAGGGCTACCGCTG 440


RESULT 9
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             3.5%;  Score 45;  DB 3;  Length 4411529;
  Best Local Similarity   53.0%;  Pred. No. 1.5;
  Matches   96;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy          1 CCCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGC 60
              |||| ||  |||||||  | |   | || | || | |||  ||||  |||     | | |
Db    1573347 CCCAGGCCGCCGGCCCCGCCGTGGCCGATCAGCCCGGCCGCCCCGCCCGCACCGCCGGCC 1573406

Qy         61 GCCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCG 120
                | || || |  || |    ||  || ||  | ||    | ||| | | | | |||
Db    1573407 TGACCGGTGGCGCCGGACCCGCCGTTGCCGCCGTTGCCGTACAGGATCCCGCCGGCCCCG 1573466

Qy        121 CCGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGG 180
              ||| |||  ||| | |     |   |||||||||     | || || | | | |
Db    1573467 CCGGCCTGCCCGGTCCCCGGCGCCCCGTCGGCGCCGTGGCCGATCAGCGGGCGCCCCAGC 1573526

Qy        181 A 181
              |
Db    1573527 A 1573527


RESULT 10
US-09-621-976-10505
; Sequence 10505, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 10505
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-10505

  Query Match             3.5%;  Score 44.6;  DB 3;  Length 433;
  Best Local Similarity   55.6%;  Pred. No. 0.011;
  Matches   75;  Conservative    5;  Mismatches   55;  Indels    0;  Gaps    0;

Qy         78 CGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCGTCCTAGCCGCTGCT 137
              ||||| | |   ||||   || | |||:|   ||| ||| || ||| ||   || ||||
Db        226 CGAAGAGGCTGTGCGAGCCCTDCCCGGSCCTCCCCAGGGCCCCCCGCCCCCTCCTCTGCC 285

Qy        138 GTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACAATGGGAGCAGCCGC 197
                || |||||  |  ||| || |   |   |: | ||  | || |  ||||||  :: |
Db        286 TGCTGCGTGGAGGCAGCCATGGGAAGGAGCCBAGGGGAGCTGGCCTGGGGGAGCGAMSCC 345

Qy        198 ACATTGCACTCCAGA 212
              |  ||  : ||| ||
Db        346 ATGTTCGMTTCCTGA 360


RESULT 11
US-09-902-540-8667
; Sequence 8667, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8667
;   LENGTH: 573
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-8667

  Query Match             3.4%;  Score 44.4;  DB 3;  Length 573;
  Best Local Similarity   49.6%;  Pred. No. 0.015;
  Matches  114;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || | ||| ||  | |||| | ||  | |||| |      ||||| | |  | |
Db        164 ACCCTTCCCGCAAGGCGGCCCAGGCCCTGCTCCCCGAAGTCGAGCGGGTGCTGGCGGAGG 223

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              |   ||  |  |||||    | | ||||| |||| |||| ||     | ||| | ||||
Db        224 CGCTGGCCGCGGGGCGGGATGTCTCGGCGGTGGCCCCGCTGGAGAAGCTGGTGGCCGCGC 283

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              | || |||   | |       | |  | ||| |       |  |||| |||  |||| |
Db        284 TGCTGGCCCACGGCGAGGCGCTCTGCCACACCGCGGCCGGACGTCTGGAGGCCGCGGAGA 343

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCC 233
```

```
Db        344 CGGCCTGGCGCCGGGCGCAGGAGCTGGAGCGCGCCGCGCACCCGACGCGC 393


RESULT 12
US-09-902-540-918
; Sequence 918, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 918
;   LENGTH: 9191
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-918

  Query Match             3.4%;  Score 44.4;  DB 3;  Length 9191;
  Best Local Similarity   49.6%;  Pred. No. 0.069;
  Matches  114;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy          4 ACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCGCC 63
              || | ||| ||  | |||| | ||  | ||||  |      ||||| | |  |  |
Db       2956 ACCCTTCCCGCAAGGCGGCCCAGGCCCTGCTCCCCGAAGTCGAGCGGGTGCTGGCGGAGG 3015

Qy         64 CTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGCCG 123
              |   ||  | |||||    | | ||||| |||| |||| ||     | || | ||||
Db       3016 CGCTGGCCGCGGGGCGGGATGTCTCGGCGGTGGCCCCGCTGGAGAAGCTGGTGGCCGCGC 3075

Qy        124 TCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGTGGCTCCGGACA 183
              | || |||   | |       | |   | ||| |      | |||| |||  |||| |
Db       3076 TGCTGGCCCACGGCGAGGCGCTCTGCCACACCGCGGCCGGACGTCTGGAGGCCGCGGAGA 3135

Qy        184 ATGGGAGCAGCCGCACATTGCACTCCAGAACAGAGACGACCCCGTCGCCC 233
              |   | |||| |  | |       |   ||  |||| ||| |
Db       3136 CGGCCTGGCGCCGGGCGCAGGAGCTGGAGCGCGCCGCGCACCCGACGCGC 3185


RESULT 13
US-10-012-231A-259
; Sequence 259, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan l.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2830P1C23
;  CURRENT APPLICATION NUMBER: US/10/012,231A
;  CURRENT FILING DATE:  2002-06-10
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 259
;   LENGTH: 4563
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 3635
;   OTHER INFORMATION: unknown base
US-10-012-231A-259

  Query Match             3.3%;  Score 43;  DB 3;  Length 4563;
  Best Local Similarity   53.2%;  Pred. No. 0.12;
  Matches   91;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy          2 CCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 61
              ||  | |   || | || ||||  ||     ||   |||| |    |||  ||  ||||||
Db        398 CCGTGGGCGGGGGCAGCGGGCTGAGGGCGCGCGGAGCCTGCGGCGGCGGCGGCGGCGGCG 457

Qy         62 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 121
                | ||   || ||||| |||| |||||  |||||  |||| |  |  |  |  |
Db        458 GCGGCGGCCCGGCGGGCGGAGCGGCGCGGGCATGGCCGCGCGCGGCCGGCGCGCCTGGCT 517

Qy        122 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGT 172
                |  | |       |   || || ||||   ||   ||   |||| ||| ||
Db        518 CAGCGTGCTGCTCGGGCTCGTCCTGGGCTTCGTGCTGGCCTCGCGGCTCGT 568


RESULT 14
US-10-015-389A-259
; Sequence 259, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan l.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2830P1C48
;  CURRENT APPLICATION NUMBER: US/10/015,389A
;  CURRENT FILING DATE:  2002-06-25
;     Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 259
;   LENGTH: 4563
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 3635
;   OTHER INFORMATION: unknown base
US-10-015-389A-259

  Query Match             3.3%;  Score 43;  DB 3;  Length 4563;
  Best Local Similarity   53.2%;  Pred. No. 0.12;
  Matches   91;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy          2 CCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 61
              ||  | |   || | || ||||  ||     ||   |||| |    |||  ||  ||||||
Db        398 CCGTGGGCGGGGGCAGCGGGCTGAGGGCGCGCGGAGCCTGCGGCGGCGGCGGCGGCGGCG 457

Qy         62 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 121
                |  ||   || ||||| |||| |||||  |||||  |||| |  |  |  |  |
```

```
Db        458 GCGGCGGCCCGGCGGGCGGAGCGGCGCGGGCATGGCCGCGCGCGGCCGGCGCGCCTGGCT 517

Qy        122 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGT 172
              |  | |      |   || || ||||   ||   ||   |||| ||| ||
Db        518 CAGCGTGCTGCTCGGGCTCGTCCTGGGCTTCGTGCTGGCCTCGCGGCTCGT 568


RESULT 15
US-10-006-768A-259
; Sequence 259, Application US/10006768A
; Patent No. 6936697
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan l.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2830P1C10
;  CURRENT APPLICATION NUMBER: US/10/006,768A
;  CURRENT FILING DATE:  2002-03-05
;  NUMBER OF SEQ ID NOS: 477
;  Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 259
;   LENGTH: 4563
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 3635
;   OTHER INFORMATION: unknown base
US-10-006-768A-259

  Query Match             3.3%;  Score 43;  DB 3;  Length 4563;
  Best Local Similarity   53.2%;  Pred. No. 0.12;
  Matches   91;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy          2 CCACGCGTCCGGCCCGCCGGCTCCGGAGCGGCTCTGCCTTCCCGAGCGCGGGACGCGGCG 61
              ||  | |   || | || ||||  ||    ||   |||| |    |||  ||  ||||||
Db        398 CCGTGGGCGGGGGCAGCGGGCTGAGGGCGCGCGGAGCCTGCGGCGGCGGCGGCGGCGGCG 457

Qy         62 CCCTGGGGGAGGAGGGCGAAGCGACGCGGCGATGGCTCCGCGGGCACTCCCGGGGTCCGC 121
              |   ||   || ||||| |||| |||||  |||||  |||| |  |  |  |  |  |
Db        458 GCGGCGGCCCGGCGGGCGGAGCGGCGCGGGCATGGCCGCGCGCGGCCGGCGCGCCTGGCT 517

Qy        122 CGTCCTAGCCGCTGCTGTCTTCGTGGGAGGCGCCGTGAGTTCGCCGCTGGT 172
              |  | |      |   || || ||||   ||   ||   |||| ||| ||
Db        518 CAGCGTGCTGCTCGGGCTCGTCCTGGGCTTCGTGCTGGCCTCGCGGCTCGT 568


Search completed: February 23, 2006, 15:12:58
Job time : 274 secs
```